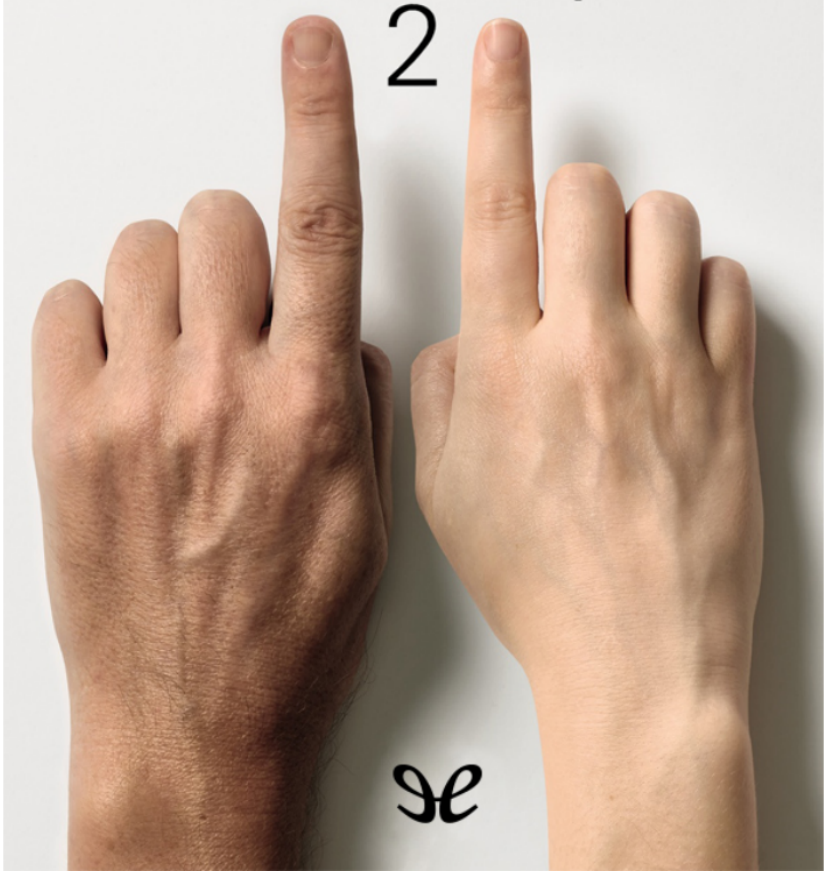
John D. Barrow

1+1 no es (siempre) 2



En 1 + 1 no es (siempre) 2, el prestigioso matemático y divulgador John D. Barrow ofrece un repaso y análisis de los fundamentos filosóficos y conceptuales de la matemática moderna. La matemática transfinita de Cantor, el teorema de Gödel o el teorema de Benford son algunos de los contenidos más conocidos que trata la obra, pero a ellos se suman otros que lo son menos, aunque igual de interesantes. Bajo la aparente simplicidad de la fórmula $1+1=2$ se ocultan numerosas y variadas posibilidades que Barrow nos descubre, no siendo la menor de ellas que modelos así de simples, como la aritmética, son tan ricos y poderosos que revelan la estructura misma del universo.

John D. Barrow

1 +1 no es (siempre) 2

Una lección de matemáticas

ePub r1.0
Titivillus 18.10.2023

Título original: *1 + 1 non fa (sempre) 2. Una lezione di matemàtica*
John D. Barrow, 2020
Traducción: Dulcinea Otero-Piñeiro
Revisión científico-técnica de la traducción: David Galadí-Enríquez, doctor en física
Retoque de cubierta: Editorial

Editor digital: Titivillus
ePub base r2.1

Índice de contenido

Cubierta

1 +1 no es (siempre) 2

Prefacio

1. 1 + 1: ¿de verdad es algo tan complejo?

2. Los dedos de las manos y de los pies: cómo empezamos a contar

3. Cambio de base: bits y bytes

4. La definición de los números

5. Sumas de conjuntos y otras cosas

6. Demostración de Whitehead y Russell de 1 + 1 = 2

7. Aritmética transfinita

8. La incompletitud de Gödel

9. Por qué son tan comunes los unos y los doses

10. ¿Qué son las matemáticas?

Índice analítico

Sobre el autor

Notas

Para Stephen, que construye,
cuenta y juega

Si la gente no cree que las matemáticas son simples es únicamente porque no se da cuenta de lo complicada que es la vida.

John von Neumann

Prefacio

Este es el último libro que escribo. No podré escribir ninguno más. Trata sobre números. Los números cuentan. Muchas personas creen que una operación como $1 + 1 = 2$ es demasiado simple para crear alguna inquietud. Sin embargo, aquí desmenuzaremos algunas de las complicaciones que plantea esta suma elemental. Analizaremos las sutilezas de sumar cosas diferentes. A lo largo del camino nos toparemos con algunos de los mayores matemáticos de los siglos XIX y XX que estudiaron esta cuestión y descubriremos cómo la abordaron y cómo explicaron el proceso de la adición. Esto nos llevará a considerar asuntos muy inusuales, como el estudio de infinitos, cómo sumarlos y los debates sobre si deben contemplarse como parte de las matemáticas. Echaremos una ojeada a una prueba simple de la famosa incompletitud de Gödel y a los acalorados debates sobre qué son en última instancia las matemáticas: «¿son un descubrimiento o son una invención nuestra?». Pero antes indagaremos en los intentos que acometieron los pueblos antiguos para desarrollar sistemas de cómputo a partir de uno, el cual, sumado a otro uno, da dos. ¿Hasta dónde consiguieron llegar? Veremos que cada sociedad primitiva desarrolló distintas formas de contar, y que casi todas las que fueron más allá de contar de uno en uno y de dos en dos acabaron coincidiendo en el cómputo por decenas y llevando la cuenta de esos grupos de decenas porque tenemos diez dedos en las manos. Por último, conoceremos algunas

propiedades absolutamente inesperadas de los números uno y dos descubiertas por primera vez de una forma casual por Simon Newcomb, aunque ahora se conocen como ley de Benford, y veremos por qué son verdad.

Los números no son lo único con lo que se puede contar. Las personas son mucho más importantes. Las personas más importantes con las que cuento yo son: Elizabeth, que ha sido mi amada esposa durante 45 años y a quien conozco desde hace al menos 55; nuestros hijos y sus parejas, David y Emma, Roger y Sophie, Louise y Stephen; y nuestros nietos, Tilly, Darcey, Mahler, Guy y Poppy.

Quisiera manifestar un agradecimiento muy especial a nuestro hijo Roger, quien durante esta etapa tan difícil ha sido una gran ayuda dando muestras de una fortaleza ilimitada. Asimismo, todo mi agradecimiento a Pino y Jo por los esfuerzos que han dedicado a materializar la idea de este libro y por haber supervisado la traducción al italiano con ayuda de algunos colegas matemáticos. Yo no podría haberlo logrado sin ellos. Que Dios los bendiga a todos.

«Cual corriente escindida por una gran roca, sé que volveremos a encontrarnos» (*A Hundred and Seventy Chinese Poems*, versión en inglés de Arthur Waley).

John D. Barrow

1. 1 + 1: ¿de verdad es algo tan complejo?

One is one and all alone
and evermore shall be so [1]

Green Grow the Rushes, O.
(Canción popular inglesa)

Durante el primer curso de la escuela de primaria todos nos encontramos con la primera fórmula de nuestra vida: 1 + 1 = 2, o sea, el asunto de este libro. Este es el primer paso para el aprendizaje de las matemáticas. Pero ¿cuánto da de sí para hablar sobre ella? ¿No es un enunciado obvio, una mera definición de lo que entendemos por «2»? Pues bien, cuando se reflexiona algo más sobre esta fórmula por primera vez, resulta que lo que dice no es tan obvio. ¿Cuánto suman una manzana y una pera? ¿Dos qué? No son dos peras ni dos manzanas. ¿Da el resultado de esta suma simplemente dos cosas como resultado? Y ¿qué son los símbolos «+» y «=»?; ¿qué significan en realidad? Si sumamos dos ondas idénticas pero con fases opuestas, de forma que las crestas de la una coincidan con los valles de la otra, lo que se obtiene es cero ondas, no dos. Si se suman un cero y un cero se obtienen dos ceros, que valen cero. Sume un infinito a un infinito y a otro infinito y obtendrá un solo infinito. Ninguna de estas sumas sigue el patrón de que la adición de uno y uno da dos. Las cosas no son tan simples como parecen. Hay que seguir ciertas reglas para que la suma de dos unidades dé dos.

Todos los sistemas de cómputo, así como los dominios científicos y técnicos creados a partir de ellos, se basan en última instancia en sistemas que empezaron sumando uno más uno. Muchos de esos

sistemas de cómputo primitivos nunca fueron más allá de sumar cada vez más y más unos entre sí, usando los dedos de las manos y de los pies como referencia para formar grupos de cinco, de diez o incluso de veinte unidades más fáciles de recordar. La naturaleza especial del número dos se nota en la lengua inglesa en la abundancia de palabras que ofrece para referirse a dos cosas: *pair, duo, brace, double, twin, duet, couple, yoke, twosome, dyad, tandem, duplet* y *twain*. Y cada una de ellas se emplea para designar dos cosas muy específicas; por ejemplo, para hablar de un par de faisanes, en inglés se usa el término *brace*; para hablar de un par de bueyes se usa la palabra *yoke* («yunta»); para referirse a un par de guantes se usa *pair*, y para aludir a una pareja de baile se emplea *couple*; pero nunca se usará *brace* para hablar de un par de zapatos, ni *couple* para nombrar un par de guantes. Esto ilustra lo indisociable que era en sus orígenes el objeto contado de la palabra empleada para indicar dos cosas.

A medida que se avanza por la escala de los números aparecen algunos términos concretos para aludir a tres cosas, como tríada, triple o trío, pero a partir de ahí estos nombres específicos desaparecen casi por completo: no hay tantos términos para referirse a cantidades como 7 u 11. Esto se aprecia en el patrón que siguen los nombres de los primeros números cardinales y ordinales en cinco lenguas europeas: castellano, inglés, francés, alemán e italiano. En ninguna de estas lenguas hay una relación entre el nombre que reciben los dos primeros números en sus versiones ordinal y cardinal, mientras que a partir del tres encontramos cierta relación, o derivación, en esos términos. Veamos los nombres de los cuatro primeros números en cada uno de esos idiomas:

Castellano: uno/primero, dos/segundo, tres/tercero, cuatro/cuarto…

Inglés: one/first, two/second, three/third, four/fourth…

Francés: un/premier, deux/second (o deuxième), trois/troisième, quatre/quatrième…

Alemán: eins/erste, zwei/ander (o zweite), drei/dritte, vier/vierte…

Italiano: uno/primo, due/secondo, tre/terzo, quattro/quarto…

En todas las lenguas indoeuropeas conocidas, los nú meros por encima de 4 son sustantivos que nunca se usan como adjetivos que puedan cambiar de forma dependiendo del tipo de objeto al que hacen referencia. Esto evidencia la antigüedad de los conceptos de unidad y duplicidad, frente a la de otras cantidades mayores. Es posible que esto guarde alguna relación con la capacidad mental humana para reconocer al instante si hay 1, 2, 3, 4 o tal vez incluso 5 objetos en un conjunto sin necesidad de enumerarlos uno por uno, ya sea mental o físicamente. Cuando son más de cinco perdemos esa capacidad y necesitamos contarlos, a menos que podamos subdividirlos en grupos más reducidos, del mismo modo que separamos los números de teléfono en grupos de dos o tres dígitos espaciados entre sí para desplazar el problema de recordarlos a la memoria a corto plazo. O bien la división puede resultar sin más del mero hecho de que tenemos cuatro dedos largos en cada mano y, en efecto, muchas civilizaciones antiguas tuvieron una unidad de longitud denominada «palmo», que en algunos lugares se corresponde con la extensión que abarcan esos cuatro dedos juntos. De hecho, el término *dígito* que usamos para aludir a un número, deriva de la palabra latina *digitus*, «dedo», mientras que una de las medidas que se sigue empleando en la actualidad, sobre todo al pedir una bebida espirituosa, es «un dedo». Un dedo de whisky equivale a llenar el vaso hasta la altura del grosor de un dedo de la mano.

En tiempos modernos hemos usado sistemas de cómputo en distintas bases, especialmente el sistema binario que se emplea con los lenguajes de computadoras. ¿Qué ocurre entonces con la fórmula 1 + 1?

Los filósofos de las matemáticas han indagado con más profundidad en el significado de nuestra sencilla fórmula, planteándose si puede demostrarse a partir de axiomas que definan los números y la operación de sumarlos entre sí. El manual clásico de 1910 titulado *Principia Mathematica*, de Bertrand Russell y Alfred North Whitehead, profesor del primero, no demuestra que 1 + 1 = 2 hasta muchos cientos de páginas después del inicio de esta obra de 2000 páginas en tres volúmenes [2]. Tras la demostración de que 1 + 1 = 2, los autores señalan de un modo encantador que ¡esta «proposición es útil en ocasiones»! Más adelante, en el capítulo

seis, veremos su demostración y la traduciremos a unos términos sencillos, comprensibles sin necesidad de recurrir a la lógica matemática. Los filósofos todavía se preguntan si esa fórmula no es más que una definición de lo que significa el número 2 o el signo +, y si se trata (junto con el resto de las matemáticas) de algo que hemos descubierto o sencillamente inventado, lo cual debatiremos en el capítulo diez.

En unos términos más ligeros, todos estamos familiarizados con los usos simples de los números pequeños, por ejemplo, los 2 puntos, o ahora 3, que se consiguen al ganar un partido de fútbol, 0 puntos por perder y el punto que logra cada equipo por empatar. El cambio a la obtención de 3 puntos por ganar es bastante incómodo desde un punto de vista numérico. Un partido con resultado de victoria para uno de los contrincantes da 3 puntos en total: 3 puntos para el equipo ganador y 0 para el equipo perdedor; pero un partido que acaba en empate otorga tan solo 2 puntos en total, uno para cada equipo. Con el sistema antiguo que otorgaba también 2 puntos al equipo vencedor, siempre se repartían 2 puntos en total con independencia del resultado final. El sistema nuevo impide prever la evolución futura de un equipo dentro de una liga de fútbol después de jugar muchos partidos. Estamos acostumbrados a la deducción simple que efectúan jugadores, entrenadores y periodistas de que el número máximo de puntos que puede reunir un equipo se corresponde con los puntos que tiene en el momento actual más todos los que obtendría si ganara todos los partidos que le quedan por jugar. Pero ¿y si queremos tener en cuenta todo el laberinto de interconexiones que crean los resultados de todos los partidos que le restan por jugar a un equipo determinado? Resulta que el viejo sistema de otorgar 2 puntos en cada partido, ya fuera con victoria de un bando o con empate entre ambos, permitía predecir las posibilidades futuras de cada equipo a partir de un instante concreto de la temporada. Sin embargo, si cada victoria otorga 3 puntos, el hecho de que un partido pueda valer 2 o 3 puntos dependiendo del resultado incrementa lo suficiente la cantidad de posibilidades como para que deje de ser matemáticamente predecible la clasificación final de cada equipo particular[3]. Sencillamente se crean demasiadas rutas futuras posibles al duplicar las opciones de cuántos puntos se asignan en

cada partido.

Veamos esta otra paradoja para ir abriendo boca. Todos estamos de acuerdo en que 1 – 1 da 0, de igual manera que –1 + 1 es 0. Además, la suma de sumas que dan 0, posiblemente infinitas sumas, siempre sigue dando 0. Pero en ese caso,

$$1 = 1 + 0 + 0 + 0 + \ldots =$$
$$= 1 + (1 - 1) + (1 - 1) + (1 - 1) + \ldots =$$
$$= 1 + 1 + (-1 + 1) + (-1 + 1) + (-1 + 1) + \ldots =$$
$$= 1 + 1 + 0 + 0 + 0 + \ldots = 2$$

De modo que $1 = 2$.

Y entonces toda la aritmética se derrumba porque, si hay una contradicción lógica dentro de un sistema axiomático, entonces se podrá utilizar para demostrar que cualquier cosa es verdadera. Un ejemplo célebre de ello ocurrió cuando un alumno respondió a esta afirmación de Bertrand Russell durante una clase retándolo a demostrar que él era el Papa de Roma. Russell respondió al instante:

> … tenemos que $2 = 1$. El conjunto que nos contiene tan solo al Papa y a mí está formado por dos miembros. Pero como $2 = 1$, consta de un solo miembro; por tanto, yo soy el Papa de Roma.

Y aquí se disparan las alarmas. Debe de haber algo equivocado en la deducción de esta ecuación. ¡Sigamos adelante!

2. Los dedos de las manos y de los pies: cómo empezamos a contar

> *«Caballero, ¿me permite una pregunta? ¿Qué haría si la Iglesia le dijera que dos más tres son diez?». «Caballero —respondió él—, lo creería, y entonces contaría del siguiente modo: uno, dos, tres, cuatro, diez». Así quedé totalmente satisfecho.*
>
> James Boswell
> (*Boswell in Holland: 1763-1764*)

La capacidad de contar es más universal que cualquier otro atributo humano, salvo el de la adquisición del lenguaje. El lenguaje parece preprogramado en el cerebro a través de un proceso evolutivo. Noam Chomsky fue el primero en proponer que este programa se configura a una edad temprana con la adquisición de pautas externas procedentes del entorno de la persona que fijan la lengua materna (la primera que se adquiere), lo que acciona varios interruptores ya existentes en el programa cerebral del lenguaje. Esto explicaría fenómenos tan extraños como que, en los primeros años de vida, parecemos saber más de lo que hemos podido aprender o nos han podido enseñar. Gran parte de ello ya está ahí esperando a ser activado.

Se ha especulado mucho sobre si ocurrirá lo mismo con la capacidad numérica. Pero, comparado con nuestras habilidades lingüísticas, el sentido numérico del ser humano es muy rudimentario, y es mucho más probable que se adquiriera por necesidades básicas de la vida o por la naturaleza de nuestra propia fisiología: cinco dedos en cada mano, diez entre las dos, veinte si

añadimos los dedos de los pies.

En todas las regiones del mundo antiguo encontramos los orígenes del cómputo representados por personas que contaban de dos en dos, o «contadores por pares». Estos pueblos seguían un sistema de cómputo que solo usaba una etiqueta para «uno» y otra etiqueta para «dos», y que construía cantidades más grandes mediante la combinación de esos dos términos. Un vestigio actual de este sistema simple en el que las cantidades superiores a dos se describieron en un principio como «más allá» o «trans» lo encontramos en los términos latinos *trans*, que significa «más allá», y *tres*, el nombre del número 3. De igual modo, en francés existen los términos *tres*, para «mucho», y *trois*, para el número 3. En particular, hay zonas de África, América del Sur y Nueva Guinea donde vemos este sistema para contar por pares ampliado a cantidades mayores con el simple recurso de la repetición, de forma que los números 3, 4 y 5 de nuestro sistema de hoy se expresan mediante «dosuno»,
«dos-dos»
y «dos-dos-uno». En general, este patrón no se sucedía de manera indefinida, tal como podría parecernos obvio en la actualidad, porque quienes contaban no consideraban cosas diferentes combinaciones tales como
«dos-dos»
o «dos-dos-uno» como cosas diferentes: solo eran la unión de las distintas agrupaciones que tenían identificadas, o sea, elementos individuales y pares. Para ir más allá en cuanto a conocimientos y capacidad de abstracción era necesario un salto conceptual mayor que solo se dio en algunas partes del mundo muy antiguo.

Sin embargo, cabría plantearse si aquellas sociedades primitivas contaban realmente o si se limitaban a realizar la suma $1 + 1 = 2$ tal como la entendemos nosotros. Usaban los nombres de los números como meros adjetivos para describir lo que veían. Los pueblos indígenas de la Columbia Británica hablantes de las lenguas tsimshiánicas usaban palabras muy diferentes para nombrar cantidades distintas de objetos diversos. Por ejemplo, usaban una palabra concreta para uno y dos cuando se trataba de una conversación general en la que no se aludía a nada específico, y otros vocablos diversos para referirse a objetos planos, curvos,

alargados, hombres, canoas o medidas:

En este ejemplo, los términos individuales no tienen ningún interés especial; lo importante es el uso adjetivado de las cantidades para describir y diferenciar cosas. Los números actuales de las lenguas europeas se apartaron hace tiempo de su forma original, y muchas de las conexiones entre números, como nombres o adjetivos, han desaparecido con el paso de los siglos porque el foco de atención se ha puesto en los símbolos de los números y las reglas para combinarlos entre sí, no en las palabras empleadas para nombrarlos. Hoy día vemos de inmediato que una fórmula como 1 + 1 = 2 solo surgiría dentro de una misma categoría de objetos, por ejemplo, los objetos curvos. Si tuviéramos una mezcla de objetos de formas distintas, entonces 1 objeto curvo más 1 objeto plano no daría como resultado 2 objetos de una sola de esas categorías. El significado de los números no es lo mismo que contar. Sin esto último no habría aritmética: nada de 1 + 1 = 2. Por tanto, se observa una transformación en el uso actual de los números frente al que les daban aquellas civilizaciones antiguas. Cuando sumamos 1 y 1, estamos añadiendo una cosa de una clase específica a otra idéntica para obtener dos cosas iguales. Si queremos especificar de qué cosas se trata, tendremos que asegurarnos de que son de la misma clase.

No hay ningún sistema basado en el reconocimiento de las relaciones que mantienen diferentes cantidades. Veamos este testimonio revelador que dejó Francis Galton 120 años atrás sobre las peculiaridades de comprar o cambiar más de un bien al mismo tiempo en un mercado. En este caso se trata de cambiar ovejas por palos de tabaco:

> Cuando se realiza un trueque hay que pagar cada oveja por separado. Así, supongamos que la tasa de cambio de una oveja asciende a dos palos de tabaco. Sin duda, un demara se quedaría perplejo si le pidiéramos dos ovejas a cambio de cuatro palos. Yo lo he hecho y me he encontrado con que el hombre separó dos de los palos y echó una mirada por encima de ellos a la oveja que estaba a punto de vender. Tras convencerse de que una de ellas estaba pagada como era debido y descubrir sorprendido que en la mano le quedaban justo los dos palos necesarios para pagar la otra oveja, lo asaltaron las dudas; el

procedimiento le pareció demasiado «fácil» para ser correcto, y volvió a examinar el primer par de palos, momento en que se le nubló la mente y, con la confusión, se paseó de una oveja a la otra; entonces interrumpió la transacción hasta tener dos palos en una mano y desprenderse de una oveja, y a continuación volver a recibir dos palos a cambio de entregar la segunda oveja.

Contar con los dedos es la base de todos los sistemas de cómputo relevantes conocidos. En algunas sociedades se añadieron todas las partes del cuerpo para contar más allá de los diez dedos. Esta ampliación es común entre los pueblos isleños del Pacífico, donde ha dado lugar a distintos totales después de sumar los diez dedos de las manos y de los pies a ambas muñecas, los codos, los hombros, el pecho, los tobillos, las rodillas y las caderas (lo que da un total de 33) u otras combinaciones de las partes del cuerpo.

Pero, volviendo a los totales más habituales para nosotros, el sistema decimal moderno, de origen indio, se basa en el número 10, de tal manera que
10×10
$= 100,$
$10 \times 10 \times 10$
$= 1000,$ etcétera. Los diez dedos de las manos motivaron esa elección. Una excepción notable la encontramos en América Central, donde los indios yuki usan un sistema en base 8 en lugar de 10, aunque también cuentan con los dedos, solo que con los huecos que hay entre ellos. Se cree que, al igual que otros pueblos de América Central y del Sur, se ataban cuerdas entre los dedos para ayudarse de ellas al contar [1].

En la actualidad todavía encontramos distintas formas de contar con los dedos de las manos. En Gran Bretaña se cuenta a partir de la mano cerrada, extendiendo los dedos uno tras otro empezando por el pulgar y terminando por el meñique antes de continuar con la otra mano si es necesaria. En Asia y Australia hay quien comienza por el dedo meñique de la mano izquierda. En Japón se empieza con la mano abierta y se van replegando los dedos uno a uno. Este método se torna explícito en el significado de los nombres que reciben los números entre el pueblo dene-dinje de América, cuya traducción es:

1 = «doblado el final» (o sea, el meñique doblado por la mitad);
2 = «doblado otra vez» (ahora se dobla el anular);
3 = «doblado el centro» (se dobla también el dedo corazón);
4 = «solo queda uno» (es decir, al doblar el dedo índice, solo queda desplegado el pulgar)
5 = «mano terminada»

No es infrecuente encontrar historias curiosas como la siguiente. Durante la Segunda Guerra Mundial, una joven india se vio en la tesitura de tener que presentar una de sus amigas orientales a un caballero inglés que acudió de visita a su casa. El problema era que aquella amiga suya era japonesa, y sería arrestada inmediatamente en caso de saberse. Así que la muchacha india dijo al visitante que era china. El inglés sospechó y sorprendió a ambas al pedir a la amiga que contara hasta cinco con los dedos. La joven india se quedó cuando menos extrañada por la curiosa petición del visitante, tal vez explicable por el consabido sentido del humor inglés, mientras la otra, imperturbable, abrió la mano y empezó a contar doblando un dedo cada vez. Entonces el caballero inglés espetó triunfal: «¡No es china, es japonesa!». Los chinos cuentan como los ingleses, partiendo de la mano cerrada y desplegando los dedos uno a uno. Los japoneses, en cambio, los van cerrando a partir de la mano abierta.

En la actualidad encontramos verdaderos contadores por pares entre los bosquimanos de África, y en Australia y América del Sur. Claramente esta práctica estuvo más extendida en el pasado. Los sistemas para contar de dos en dos dan lugar a vocablos para nombrar los números siguiendo un patrón que, traducido a nuestros símbolos, sería: 1, 1 + 1 = 2, 2 + 1, 2 + 2, 2 + 2 + 1, 2 + 2 + 2, y así sucesivamente hasta llegar a 2 + 2 + 2 + 2 + 2, si se basan en contar con los dedos hasta diez. Nótese que la palabra dos es fundamental para formar el nombre del resto de los números. Además, los contadores por pares evidencian que no siempre se empezó a contar a partir de los dedos. Cuando se cuenta con los dedos se tiene una palabra para nombrar el número seis, que está formada por las palabras cinco + 1, y no dos + dos + dos, que es lo que resultaría de un sistema que cuenta de dos en dos. Este sistema existía en Sumeria en el año 3000 a. C., pero con posterioridad quedó desbancado por sistemas más eficaces. El dos

como base es demasiado bajo para resultar práctico al contar cantidades grandes. Algunos sistemas para contar por pares usan tres términos: uno, dos y muchos [2].

Lo que se observa en estos antiguos sistemas simples para contar por pares son las primeras posibilidades de expresar una fórmula como 1 + 1 = 2, que constituye el primer paso para ascender por una escala de progresión que parte de 1. Pero no se trata de un sistema aritmético como lo entendemos hoy, sino de un mero procedimiento para etiquetar cantidades mediante una serie de palabras en lugar de asignar un término diferente a cada cantidad, lo que sería equivalente a contar con los vocablos uno, dos, tres, cuatro, cinco, etc., pero sin los símbolos aritméticos ni las reglas para combinar los números entre sí.

Por último, debemos retomar nuestra fórmula clásica 1 + 1 = 2 para plantearnos de dónde salieron esos símbolos. El 1 es una marca antigua de una cantidad individual, usada en muchos sistemas de cómputo porque no es más que la representación de un dedo. El 2 es un símbolo que deriva de los primeros sistemas de cómputo indios. Evolucionó a través de múltiples variantes, y al principio se representaba mediante dos barras horizontales (o unos) =, que más tarde se fundieron entre sí para adoptar una forma de z que poco a poco se fue convirtiendo en un 2 más redondeado, pero existen muchas variantes y distorsiones regionales debidas a distintas grafías y anotaciones rápidas a mano. Los signos restantes de esta ecuación son + y =. El signo de la adición es una abreviatura del término latino *et*, que significa 'y', para escribir rápido a mano. El símbolo que usamos ahora es una cruz griega, más que una cruz de Malta o una cruz latina, que tienen una forma un tanto diferente. Hay una relación clara aquí con el simbolismo del cristianismo temprano.

El signo de la sustracción, –, surgió como abreviatura de m y de m–, ambos utilizados con el significado «menos», y parece que fue usado por comerciantes para indicar el peso de una mercancía después de restarle el peso del embalaje. El resultado se denominaba *minus* o *tara*, un término que sigue usándose hoy en día en el transporte de bienes comerciales. Otras variantes del signo – fueron – – o…, hasta llegar al signo que se usa en la actualidad para la división: ÷. La notación alternativa para la división,

introducida en 1651, son los dos puntos, :, de manera que A dividido entre B se indica mediante A : B.

El signo de igual, =, lo introdujo en 1557 el matemático inglés Robert Recorde en su libro sobre álgebra titulado *The Whetstone of Whitte* ['La piedra de afilar de Witte']. Lo representó con un par de líneas paralelas porque, según escribe Recorde, «*noe two thynges can be moare equalle* [3]». Su símbolo era más largo que el que se emplea hoy, algo como ==, pero esta convención tardó bastante en adoptarse de manera generalizada. Algunos matemáticos usaban incluso los símbolos || o // o)=(, en lugar de =, a pesar de la presión para contar con formatos de símbolos estandarizados debido al empleo de tipos móviles en la imprenta a partir de 1456. Probablemente se corresponda con una forma rápida de escribir el diptongo æ, que era la abreviatura de æ*quales*, el término latino para «igual». El signo = de Recorde no volvió a aparecer impreso hasta 1618. Ochenta años después el filósofo francés René Descartes todavía usaba un signo alfa, α, de proporcionalidad para denotar igualdad y recurría al símbolo = para indicar más menos, ±.

Los intereses de Recorde eran muy amplios. Su libro titulado *The Urinal of Physick* [4] [«El orinal de la medicina»], de 1547 y escrito en inglés en vez de latín, es una historia de los estudios y diagnósticos médicos realizados con orina. Cuando se publicó la edición definitiva en 1679, el título elegido fue *The Judgement of Urine* [«El dictamen de la orina»]. Recorde se graduó en Oxford en 1531 y, por tanto, estaba autorizado para ejercer la medicina. Sus intereses eran, sin duda alguna, muy amplios.

La cruz de san Andrés que usamos como signo de multiplicación, ×, aparece por primera vez en 1618 en una obra de John Napier, el inventor de los logaritmos [5]. Hoy en día todo el mundo usa los símbolos +, –, × y ÷ , más que ningún otro lenguaje o alfabeto [6].

En la tabla 2.1 se muestra la forma en que se habría escrito la fórmula 1 + 1 = 2 en diversos sistemas de cómputo antiguos. Nótese que no disponían de los signos + y =, de modo que hemos usado palabras para indicar las sumas.

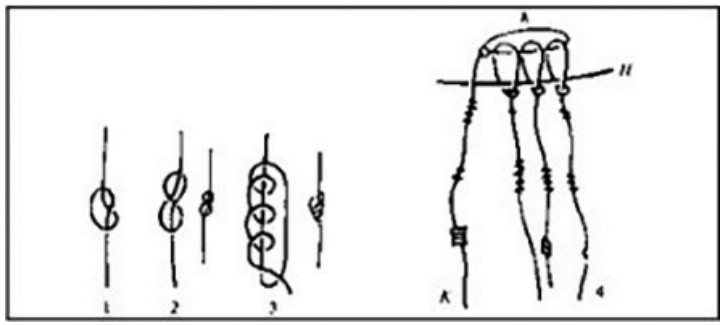


Figura 2.1. Nudos en un quipu peruano.

3. Cambio de base: bits y bytes

1 + 1 = 10

Aritmética binaria

La aritmética necesita unos fundamentos para no tener que tratar cada número como algo independiente. Estos fundamentos constituyen lo que se denomina la base de una aritmética. La base que solemos usar nosotros es 10, y deriva de cuando la humanidad usaba los dedos para contar. Algunas culturas antiguas usaron el 5 o el 20 como base y hasta el 60 como base adicional, como en el caso de los babilonios. Quedan vestigios de estos sistemas antiguos en los nombres de algunos números y de sistemas de cómputo alternativos que se usan con fines muy particulares. En inglés existe el término *score* (= 20), que deriva de la incisión que se practicaba en los palos de contabilidad para indicar veintenas, –, de ahí que en la lengua inglesa actual tenga el doble significado de «muesca» y «veinte». Asimismo, encontramos particiones de sesenta unidades en los segundos y los minutos en que se dividen las horas, así como en la medición de ángulos circulares de
6×60
= 360 grados. La lengua francesa introdujo un vocablo nuevo para nombrar el número 60 (*soixante*) en el cómputo de unidades de diez, lo que indica un viejo cambio en ese punto tras el empleo de palabras compuestas que solo llegaban hasta 50.

Hoy en día todas las civilizaciones desarrolladas usan el sistema decimal, de base 10, que se inventó en India y se extendió por la prevalencia de las lenguas indoeuropeas que fomentaron el comercio y otras formas de interacción entre India, los países árabes

y Europa. Cuando una nación con un sistema decimal se topaba con otra que usaba un sistema más enrevesado (como los numerales romanos), los comerciantes de esta última tendían a adoptar el sistema de la primera reconociendo con rapidez su sencillez para la contabilidad. Esto es lo que sucedió cuando el sistema indio en base 10 se topó con los numerales romanos que usaban los comerciantes en Italia.

En el sistema en base 10, el número 44 significa
(4×10)
$+$
(4×1);
si vemos el número 969 significa $(9 \times 100) + (6 \times 10) +$
(9×1).
Si usáramos un sistema en base 5, el número 44 significaría (4×5) $+ (4 \times 1)$. Existen tres tipos de sistemas numéricos con una base general que etiquetaremos como B. El sistema aditivo que se usaba en la Antigüedad egipcia y griega emplea símbolos diferentes para indicar las cantidades derivadas de esa base en la siguiente lista:

$$1, 2, 3\ldots, B-1, B, 2B, 3B\ldots, B(B-1);$$
$$B2, 2B2, 3B2\ldots, B2(B-1)\ldots, \text{etc.}$$

El sistema multiplicativo, usado, por ejemplo, por los chinos, requiere muchos menos símbolos, únicamente los de los números

$$1, 2, 3\ldots, B-1, B, B2, B3\ldots, \text{etc.}$$

El tercer tipo, el sistema posicional que heredamos de la cultura india, es incluso más económico. Introduce la idea del valor posicional, de tal modo que la posición en la que se encuentra cada símbolo numérico tiene un significado. Por tanto, para los romanos, el número 111 significaba «tres», mientras que para nosotros significa «ciento once». Este sistema requiere algún signo para indicar que no hay nada en una de las posiciones, de manera que 1 1 signifique ciento uno y se pueda diferenciar del número 11, que es once. Los babilonios, mayas e indios se dieron cuenta de que los sistemas posicionales necesitan un símbolo para indicar la posición vacía en 1 1 porque, si no, en cualquier anotación manual descuidada dejaría de apreciarse el hueco entre ambos unos. Así fue

como se inventó el símbolo del cero [1] para llenar ese vacío posicional, de forma que el número ciento uno se indicaba mediante 101 en el sistema indio. De modo que, con la incorporación del cero, los únicos símbolos necesarios para este sistema fueron ahora

$$0, 1, 2, 3\ldots, B - 1.$$

A partir de ahí cualquier número se puede escribir como

$$\mathrm{N} = a_n \mathrm{B}n + a_{n-1} \mathrm{B}_{n-1} + a_{n-2} \mathrm{B}_{n-2} + \ldots + a_2 \mathrm{B}2 + a_1 \mathrm{B} + a_0$$

Y lo escribimos tan solo en términos de una sucesión de aes como en

$$N = a_n\, a_{n-1}\, a_{n-2} \ldots\, a_2\, a_1\, a_0.$$

Veamos un ejemplo de este sistema. Consideremos el número 1204, que leeríamos como mil dos cientos cuatro. Como nuestra base B es 10, el índice que da el número de entradas es $n = 3$, y $N = 1204$ se expresa como (nótese que $100 = 1$)

$$1204 = \mathbf{1} \times 10_3 + \mathbf{2} \times 10_2 + \mathbf{0} \times 10 + \mathbf{4} \times 1,$$

donde las aes aparecen resaltadas en negrita, aunque ese número se escribe sencillamente como la sucesión de aes

$$1204.$$

Las potencias de 10 por las que se multiplica cada término se indican de manera implícita en las posiciones que ocupa cada número, las cuales cobran ahora gran relevancia: 1204 no es lo mismo que 1024 o 4021.

Los primeros sistemas aritméticos fueron los desarrollados por los sumerios, en el sur de Babilonia, y por los egipcios. El sistema egipcio ilustra con claridad cómo funcionaba la adición. Los símbolos de sus jeroglíficos, que ya hacían referencia a potencias de 10, eran los siguientes (entre paréntesis se indican los valores que leeríamos nosotros). Usaban la barra 1 para formar las primeras cantidades

1 (1), 11 (2), 111 (3), 1111 (4),
11111 (5)…, 111111111 (9),

y a partir de ahí introducían una cuña ∩ (para el 10) y el símbolo ℘ para el 100, seguidos de otros símbolos adicionales para 1000, 10 000, 100 000 y 1 000 000. Así, por ejemplo, 84 se escribiría ∩∩∩∩∩∩∩∩1111, y 67 sería ∩∩∩∩∩∩1111111. Una suma como 1 + 1 = 2 se convertiría sencillamente en 1 y 1, o sea, 11. Para sumar 84 y 67 se fundirían los dos bloques anteriores

∩∩∩∩∩∩∩∩1111,
∩∩∩∩∩∩1111111,

lo que daría ∩∩∩∩∩∩∩∩∩∩∩∩∩∩∩∩11111111111 (14 cuñas y 11 barras), lo que equivaldría a ∩∩∩∩∩∩∩ ∩∩∩∩∩∩∩∩1 (una cuña más en lugar de 10 barras y una barra 1) y, por último, ℘∩∩∩∩∩1, o sea, 151.

Aunque el sistema decimal, en base 10, es ubicuo en la actualidad en toda la actividad científica, lo cierto es que usamos con mucha frecuencia un sistema en otra base con unos fines muy específicos. Se trata del sistema binario, en base 2, utilizado ampliamente en ciencias de la computación. Los dígitos 0 y 1 que lo conforman se denominan bits (abreviatura de binary digits, «dígitos binarios») [2] .

En esta base, formada tan solo por los dos símbolos 0 y 1, tenemos las cuatro sumas básicas:

En este último caso de 1 + 1, la suma de los dígitos da un resultado mayor que 1, que es la cifra más alta disponible. Para realizar esa suma procederemos de la misma manera que para calcular 9 + 4 en el sistema de base 10, donde se alude a las potencias 1, 10, 100… de 10, y los dígitos disponibles van de 0 a 9; la suma anterior se convierte, por tanto, en 13 = 1 × 10 + 3, es decir, 1 decena y 3 unidades. En cambio, al calcular en base 2, donde los dígitos disponibles son solo 0 y 1, se alude a las potencias 1, 2, 4… de 2, por lo que 1 + 1 se convierte en 10, ya que su descomposición sería 1 × 2 + 0. Esto permite saber que 1 + 1 no

da 2, por el simple hecho de que 2 ha dejado de ser una cifra aquí y se representa como 10.

De igual manera, 3, 4, 5, 6, 7… se escriben 11, 100, 101, 110, 111, y así sucesivamente.

En aritmética, para cualquier base B mayor que 2, sigue valiendo la fórmula

$$1 + 1 = 2.$$

Por tanto, y resumiendo, para diferentes bases tenemos que

$$1 + 1 = 10 \text{ para } B = 2,$$
$$1 + 1 = 2 \text{ para } B > 2.$$

Para la multiplicación también se dan cuatro casos:

$$0 \times 0 = 0,$$
$$1 \times 0 = 0,$$
$$0 \times 1 = 0,$$
$$1 \times 1 = 1.$$

Sería bastante laborioso realizar una división o multiplicación larga. Probemos con 26×12, que sería $11\,010 \times 1\,100$. El resultado sería $312 = 100\,111\,000$, y los ordenadores lo calculan en un instante. Cuando pulsamos el número 3 en un teclado de ordenador, la máquina registra el número 11 en binario. El número 240 se registrará como 11 110 000, que es la forma binaria del desarrollo decimal, $128 + 64 + 32 + 16 = 240$.

El uso moderno del bit está muy relacionado con la capacidad básica para distinguir entre dos posibilidades (o una posibilidad binaria): SÍ/NO, ENCENDIDO/APAGADO o VERDADERO/FALSO en una computadora digital. Verdadero suele corresponderse con 1, y falso, con 0. Las dos posiciones de un interruptor eléctrico lo facilitan si ambas posibilidades se distinguen con diferentes voltajes.

La invención de esta aritmética binaria moderna se suele atribuir al filósofo y matemático alemán Gottfried Leibniz, coetáneo y rival de Isaac Newton, por una explicación que expuso en 1679 [3], aunque cien años antes ya la había concebido y usado el ignorado

científico inglés Thomas Hariot[4]. Leibniz tuvo la gran idea de convertir todos los enunciados verbales en una expresión aritmética, de modo que las consecuencias de determinados supuestos pudieran inferirse de un modo sistemático. Tuvo la ambiciosa ilusión de que con ello pudieran resolverse todas las disputas políticas, religiosas y científicas con la mera aplicación de las reglas de la aritmética a los supuestos de partida previamente transformados de palabras en números. Asimismo, imaginó que podría lograrse de forma mecánica mediante lo que hoy denominaríamos una computadora[5]. El sistema binario de Leibniz era idéntico al que usamos en el momento presente:

$$\begin{aligned} &0\,0\,0\,1 \text{ tiene el valor } 2_0 = 1,\\ &0\,0\,1\,0 \text{ tiene el valor } 2_1 = 2,\\ &0\,1\,0\,0 \text{ tiene el valor } 2_2 = 4,\\ &1\,0\,0\,0 \text{ tiene el valor } 2_3 = 8, \end{aligned}$$

y así sucesivamente.

Como sinólogo entusiasta que era, Leibniz se inspiró en los hexagramas del libro chino *I Ching*, que se correspondían con los números binarios del 0 al 111 111.

Encontramos una aplicación importante de un código binario en el sistema Braille de lectoescritura táctil, desarrollado por Louis Braille entre 1829 y 1837, y utilizado desde entonces por las personas invidentes (aunque en la actualidad solo lo usa un 1% de las personas ciegas registradas en Reino Unido). Es la primera forma de escritura binaria que se desarrolló en la historia. La escritura habitual de las lenguas indoeuropeas de hoy usa las 26 letras del alfabeto fenicio, pero el sistema Braille utiliza solo dos: puntos lisos y puntos en relieve.

Otro sistema de comunicación y de enumeración que solo usa dos símbolos (puntos y rayas) es el código Morse, inventado por el artista estadounidense Samuel Morse en 1837. Este sistema codifica el alfabeto fenicio, los signos de puntuación, algunos símbolos adicionales y los diez numerales básicos (0, 1…, 9) mediante patrones de puntos y rayas. También se han creado alfabetos Morse para idiomas diferentes del inglés (véase la figura 3.1). El primer sistema de Morse se enviaba mediante pulsos eléctricos, pero en

1844 el empleo de corrientes eléctricas permitió imprimir mensajes con hendiduras sobre el papel para que pudieran leerlos de manera táctil personas ciegas.

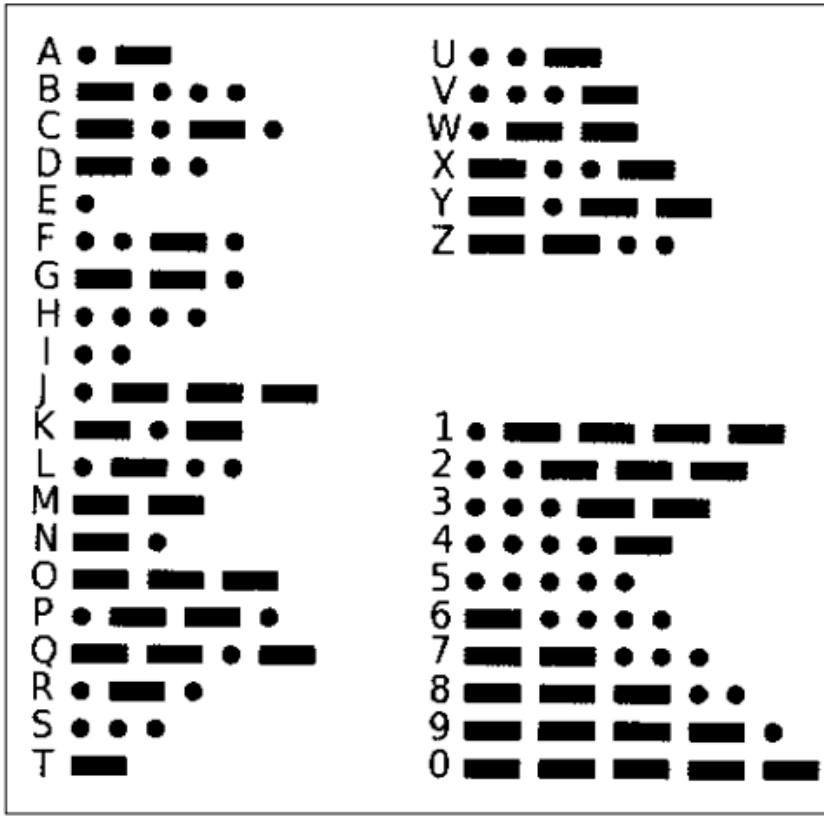


Figura 3.1. Código Morse internacional.

Algo difícil de realizar con aritmética binaria son las fracciones. En nuestro sistema decimal estamos familiarizados con fracciones como 1/3, 3/10 o 2/10. Esta última se puede simplificar porque su denominador es divisible entre 2, equivalente a 2×5, lo que permite reducir la fracción a $2/(2 \times 5) = 1/5$.

Expresar estas fracciones en base 2 es sencillo: 1/3 y las otras se convierten en 1/11, 11/111, 11/1010, 10/1010, que se simplifica en 1/101. Pero ciertamente estamos menos familiarizados con estas representaciones.

Además, existe un desarrollo de las fracciones en base 2 similar al que denominamos decimal. En este último caso utilizamos potencias de 10 y sus inversos, es decir, 1/10, 1/100, etcétera. Así, una fracción como 2/10, es decir, 1/5, se escribe 2 × 1/10 y equivale, como sabemos, a 0.2. De manera similar 3/10 corresponde a 3 × 1/10, es decir, 0.3. En cuanto a 1/3, la división habitual de 1 entre 3 se representa como 0.333 3… es decir, como 3 × 1/10 + 3 × 1/100 + 3 × 1/1000 +…, etc., o sea, en forma decimal infinita y periódica.

Lo mismo sucede con las fracciones en base 2, solo que en este caso aludiremos a las potencias de 2, no de 10. De modo que

- 1/5, o 0.2 en base 10, se escribe 1/8 + 1/16 + 1/128 + 1/256 +…, en detalle 0 × 1/2 + 0 × 1/4 + 1 × 1/8 + 1 × 1/16 + 0 × 1/32 + 0 × 1/64 + 1 × 1/128 + 1 × 1/256 +…, lo que da la expresión 0.001 100 11…,
- 1/3, o 0.333… en base 10, se convierte en 1/4 + 1/16 + 1/64 +…, es decir, 0 × 1/2 + 1 × 1/4 + 0 × 1/8 + 1 × 1/16 + 0 × 1/32 + 1 × 1/64 + 0 × 1/128 + 1 × 1/256 +…, o sea, todavía 0.010 101 01…,
- 1/2 es sencillamente 0.1, ya que es 1 × 1/2.

Todo esto requiere cálculos más complejos o menos habituales, pero, aun así, da lugar a expresiones que son, o bien finitas, o bien infinitas y periódicas.

En años recientes se ha incorporado al vocabulario de la ciencia el término cúbit, inventado en 1995 por Benjamin Schumacher. Si el bit es la unidad fundamental (la más pequeña) de información en computación, el cúbit constituye la unidad más pequeña de información cuántica. El dígito binario, o bit, igual a 0 o 1, se usa para aludir a una unidad básica de información, pero se creó teniendo en cuenta una computadora clásica, no cuántica. En mecánica cuántica, el contenido de información de un estado del mundo puede ser una combinación lineal ponderada de 0 y 1, un estado mixto, en lugar de 0 o 1 [6]. Este es el origen de la célebre paradoja del gato de Erwin Schrödinger.

Él imaginó un gato en una habitación donde una transición cuántica a un nivel subatómico pudiera liberar o no un veneno mortal. Si se piensa en el resultado de acuerdo con la física clásica,

el gato estará o vivo o muerto. Pero de acuerdo con la mecánica cuántica, podrá encontrarse en una cantidad ilimitada de estados mixtos en los que sea ¡una mezcla de gato vivo y muerto al mismo tiempo! Este rasgo de la teoría cuántica es equivalente a la noción de que cuando un estado cambia a otro no se limita a seguir una sola ruta evolutiva, como ocurriría con una piedra que sigue una trayectoria parabólica en el aire, sino que sigue todas las rutas que podría tomar. Algunas de ellas son como las crestas de una onda, mientras que otras son como los valles. Cuando se suman, la mayoría de ellas se cancela dejando como más probable la trayectoria parabólica clásica que nos es familiar. Son olas de información, pero se parecen más a una ola de crímenes (si se produce una ola de crímenes en tu vecindario es más probable que encuentres un crimen cometido en él) que a una ola del mar. Las personas que investigan sobre el potencial y la fabricación de computadoras cuánticas esperan poder crear un tipo nuevo de computadora superveloz basada en estas características cuánticas para realizar muchos cálculos distintos de idéntica complejidad al mismo tiempo antes de recombinarlos en una sola respuesta. Asimismo, confían en que algunos cálculos imposibles de realizar en la actualidad por la cantidad de tiempo que requerirían puedan quedar reducidos a tareas computacionales rápidas y practicables que permitan transformar lo incalculable en calculable. Puede que estos sean algunos de los descubrimientos del futuro.

4. La definición de los números

A veces me he preguntado si un cerebro como el de Von Neumann no es señal de una especie superior de hombre.

Hans Bethe [1].

Los matemáticos del siglo XIX empezaron a preocuparse más en serio por los fundamentos de las matemáticas y la necesidad de demostrar todas aquellas afirmaciones que se habían limitado a aceptar como verdaderas por el mero hecho de que parecían obvias. Aunque se trataba de unos recelos muy pedantes en numerosos aspectos, se basaban en el temor a que pudiera haber premisas falsas en alguna aplicación de la aritmética capaces de desmoronar todo el entramado por contradicciones internas si, tal como se dijo en el capítulo uno, una premisa falsa de la aritmética permitía deducir cualquier cosa como, por ejemplo, que ¡1 + 1 fuera igual a cualquier resultado deseado! Esta es una regla antigua de la lógica según la cual *ex falso quod libet*, es decir, cualquier deducción puede demostrarse verdadera a partir de una premisa falsa.

Una de las recetas defendidas para evitar este posible derrumbe de las matemáticas fue el formalismo. La idea es simple. En el caso de la aritmética, se especifican las reglas de «juego» y los participantes (los números enteros positivos) y, a partir de un punto de partida coherente, se extraen todas las verdades aritméticas posibles derivadas de la aplicación de las reglas para los números. De este modo aparecerán ante nosotros todas las verdades de la aritmética en forma de deducciones válidas derivadas de la aplicación de las reglas del juego. La misma idea podría aplicarse a un juego basado en reglas, como el ajedrez (cuyas piezas no tienen

ningún significado fuera de él). A partir de la posición inicial habitual hay ciertas configuraciones de piezas que no podrán darse jamás, y otras que se podrían comprobar para asegurarse de que en efecto derivan de una cantidad finita de movimientos desde la posición de partida[2]. Parecía una receta irrefutable, y funcionó con la geometría de Euclides (con algunas reglas adicionales que se añadieron para mejorarla[3]), pero, como veremos en el capítulo ocho, para sorpresa de todos se reveló como una falsa esperanza. Había enunciados numéricos que no se podían demostrar verdaderos ni falsos con las reglas de la aritmética. Es como ver una disposición de las piezas de ajedrez sobre el tablero y descubrir que es imposible saber si podría llegarse a ella a partir de las posiciones de partida. Esto se podría hacer demostrando simplemente que no hay ninguna configuración previa que dé lugar a las posiciones actuales sobre el tablero y, en tal caso, tendríamos lo que se denominan configuraciones «Jardín del Edén».

El matemático que contribuyó de un modo más patente a la idea de definir los números naturales de forma rigurosa fue el italiano Giuseppe Peano[4] en 1889, si bien el planteamiento de base lo introdujo en 1884 Gottlob Frege en su obra *Die Grundlagen der Arithmetik*[5], aunque esta obra no trascendió ampliamente hasta finales del siglo XIX. Richard Dedekind también había propuesto axiomas similares en 1888 en su obra *Was sind und was sollen die Zahlen?*[6], pero más tarde Peano simplificó su propuesta.

Él introdujo cinco reglas que ahora se conocen como los postulados de Peano para definir los números enteros y la aritmética. La idea clave fue el empleo de una instrucción llamada sucesor que desplaza cualquier número hasta su sucesor, por ejemplo, el 1 al 2, el 2 al 3, etcétera. Con ello demostró que se puede crear un conjunto infinito de cantidades (0, 1, 2, 3…) a partir de un conjunto finito de reglas. Tomando el 0 como número de partida (Peano usó en realidad el 1, pero no hay ninguna diferencia, ya que 1 es el sucesor de 0), entonces los cinco postulados de Peano para definir los números naturales no negativos son:

1. El cero es un número natural.
2. Todo número natural tiene un sucesor que es uno de los números naturales.

3. El cero no es sucesor de ningún otro número natural.
4. Si dos números naturales tienen el mismo sucesor entonces son el mismo número.
5. Si un conjunto contiene el cero y el sucesor de todos los números, entonces contiene todos los números naturales. Esto se denomina principio de inducción.

Todos estos axiomas son independientes y, por tanto, necesarios.

Los objetos primarios postulados aquí (no demostrados) son la cantidad cero, 0, y la operación del sucesor, que asocia cada número al que le sigue inmediatamente e indicaremos mediante S. Por tanto, $S(0)$ se identifica con 1, $S(1)$, con 2, y así sucesivamente.

El principio de inducción también permite introducir las operaciones aritméticas habituales de la suma y la multiplicación. Para cada número n, definiremos la suma y el producto de n por un genérico m de manera recursiva, suponiendo primero que m es 0 y pasando después de m a $S(m)$.

Aritmética: Consideremos en primer lugar la adición. Para ello escribiremos

$$n + 0 = n \text{ (por definición)}$$
$$n + S(m) = S(n + m) \text{ (según el postulado 2).}$$

Por tanto, de esta definición se obtiene que $n + 1 = n + S(0) = S(n + 0) = S(n)$, aplicando los postulados 1 y 2, por lo que el funcionamiento del sucesor equivale a sumar 1. Si ponemos ahora que $n = 1$, entonces vemos que con este sistema de reglas se demuestra que $1 + 1 = S(1) = 2$.

A partir de ahí se puede seguir avanzando de manera recursiva; es decir, aplicando la misma regla una y otra vez para obtener, por ejemplo, que $n + 2 = n + S(1) = S(n + 1) = S(S(n))$, donde la aplicación del sucesor dos veces equivale a sumar 2.

En general, para cualquier número natural m, la suma de n y m equivale a aplicar S una cantidad m de veces: $n + m = S(\ldots S(n))$, donde los puntos suspensivos indican que S interviene m veces.

Consideremos ahora la multiplicación. En general se tiende a pensar que es una forma repetida de adición: $n \times m$ equivale a

sumar el elemento n consigo mismo m veces. Así, cuando se pasa de la multiplicación de $n \times m$ a $n \times S(m)$, el número de veces que se suma n aumenta en uno. Confiando entonces en el principio de inducción, pondremos

$$n \times 0 = 0,$$
$$n \times S(m) = n \times m + n.$$

El esquema de Peano tiene un problema, porque no solo define los números naturales. Sus cinco axiomas se pueden usar para definir cualquier progresión infinita de cantidades. Supongamos que reemplazamos los números naturales por los números pares, de tal forma que la progresión sea

0, 2, 4, 6, 8, 10…

Claramente, la nueva secuencia satisface los postulados de Peano: parte del 0 y luego avanza de manera sucesiva de un término a otro sumando 2 cada vez. De modo que ahora $S(n) = n + 2$.

Lo mismo puede decirse de los números impares 1, 3, 5, 7…, que parten del 1 y se suceden volviendo a sumarles 2: $S(n) = n + 2$ también esta vez.

Pero la cosa empeora aún más porque hay una cantidad infinita de progresiones que cumplen los cinco postulados de Peano. Supongamos que tenemos una progresión infinita de términos:

$$N_0, N_1, N_2, N_3, N_4\ldots$$

A cada término de esa serie se puede llegar partiendo de N_0 mediante una cantidad finita de pasos. Como este término carece de sucesor, desempeña el papel de 0, mientras que el sucesor de cualquier término N_n es N_{n+1}. De modo que solo necesitamos un término de partida y una progresión infinita sin términos repetidos que permita obtener cada uno de sus términos desde el inicio mediante la aplicación repetida de una regla una cantidad finita de veces. Las series que obedecen las cinco reglas de Peano se denominan progresiones. Pero en realidad todas las progresiones, tal como las hemos descrito, se parecen a los números naturales,

con la salvedad de que en el caso de los naturales cabe la licencia de identificar el elemento genérico N_n con su índice n. Como se suele decir en matemáticas con el nombre técnico adecuado, todas las progresiones son isomorfas a la de los números naturales, en tanto que, en esencia y más allá de la apariencia, la reproducen.

Reiteramos, sin embargo, que el esquema de Peano no se aplica únicamente a los números naturales, aunque esta sea una de las progresiones definidas por él.

La primera vez que se propusieron los axiomas de Peano, algunos especialistas en lógica, como Bertrand Russell, pensaron que definían en su totalidad lo que se entiende por número natural, aunque pasaban claramente por alto la cuestión de la no unicidad que acabamos de discutir; pero otros, como el matemático francés Henri Poincaré, se mostraron escépticos. Poincaré señaló que ese caso solo se daría si los axiomas se demostraban coherentes. Porque si permitían deducir, por ejemplo, que $1 = 2$, entonces serían incoherentes e inútiles para definir nada.

Existe una versión más débil de la aritmética de Peano denominada PA (de *Peano Arithmetic*) y que se enuncia con lo que se conoce como lógica de primer orden (el enfoque lógico más adecuado para las matemáticas). Es más simple y fluida, pero no incorpora en su totalidad el poder del principio de inducción.

Fue mucho después, en 1936, cuando el joven matemático alemán Gerhard Gentzen (fallecido en un campo de trabajo soviético de Praga en 1945) demostró que los axiomas de la PA son, en efecto, coherentes y no dan lugar a contracciones. Casi todos los matemáticos de hoy aceptarían la validez del esquema axiomático de Peano, el cual lo infiere todo a partir de tres únicas nociones: la de número natural, que hay que definir, y las de 0 y sucesor, que vienen dadas.

Pero hay algunos matemáticos que no lo hacen. Son los finitistas, quienes no aceptan ningún razonamiento que maneje cantidades infinitas y pasos deductivos. Para ellos, aceptar los axiomas de Peano es como creer en un conjunto infinito de números naturales. La mayoría de los especialistas en matemáticas la consideran una postura un tanto pedante. Al fin y al cabo, tal como veremos con posterioridad, los números naturales son una infinidad contable (la infinidad más pequeña posible) del tipo que Aristóteles

admitía y llamaba infinito en potencia. Es un infinito que no se puede alcanzar ni puede influir en nosotros como podría hacerlo un infinito en acto (como una temperatura o densidad infinitas en algún lugar del universo), lo que Aristóteles había prohibido junto con el vacío local, porque eliminaría toda resistencia y permitiría que el movimiento alcanzara una velocidad infinita en un tiempo finito.

En la definición y aplicación de la función sucesor se reconocen las intuiciones de las culturas antiguas que desarrollaron sistemas de cómputo basados en sumar 1 a las cantidades ya existentes y, dependiendo de su sofisticación, en la repetición de ese procedimiento una y otra vez.

La función sucesor parece ser un concepto natural para la mente humana. En cierto sentido parece una consecuencia de la existencia de la flecha del tiempo y de la forma en que experimentamos el orden en que se producen las causas y los efectos. Distinguimos entre pasado y futuro, y el futuro avanza hacia delante con una sucesión de causas y efectos. El relato corto de Ted Chiang titulado «La historia de tu vida», y la película basada en él de Denis Villeneuve titulada *La llegada* (*Arrival*, 2016), presentan un contraste interesante. Doce naves espaciales aterrizan en el planeta Tierra con intenciones misteriosas (pero totalmente pacíficas). Entonces se recurre a una lingüista, encarnada por Amy Adams, para intentar descifrar la lengua en la que esos seres se comunican, que es extraña y atemporal. Transmiten sus mensajes con tintas de colores inyectadas en un líquido, donde crean enroscados y exóticos patrones como los que se forman al derramar un poco de tinta en agua. La lingüista descifra partes de ese lenguaje y acaba reparando en que las mentes que hay detrás de esta inteligencia avanzada no tienen la noción de tiempo. Los complejos patrones cuentan toda su historia de una sola vez. No es como nuestros relatos secuenciales de acontecimientos. De hecho, es una narración circular en lugar de lineal. Al final se descubrirá que los mensajes que envían incluyen nuestro futuro y nuestro pasado, pero son lo bastante simples como para que los humanos puedan descifrarlos en parte. La historia contiene muchas referencias, pero el aspecto lingüístico se basa en una hipótesis de 1956 del antropólogo Edward Sapir y su alumno, el lingüista Benjamin Whorf [7].

La hipótesis de Sapir-Whorf sostiene que la lengua no es tan solo una vía de comunicación, sino que también determina qué nociones somos capaces de concebir y de comunicar (aunque esta no es una teoría muy popular entre los lingüistas actuales). No es algo absoluto. El ejemplo clásico de Whorf consistía en decir que los esquimales tienen muchas palabras para describir las distintas variedades de nieve, mientras que las personas que viven en zonas más templadas solo tienen una o dos (si es que cuentan con alguna). La realidad física deja su impronta en la lengua y, entonces, la lengua impone ciertas limitaciones al pensamiento. Los filósofos de la antigua Grecia fueron grandes pensadores, pero contaron con la ayuda de que el griego clásico permitía establecer las finísimas gradaciones necesarias para los sutiles pensamientos, debates y textos filosóficos. En mi opinión, la película *La llegada* es más profunda que otras películas convencionales modernas. Es una cinta de ciencia-ficción sobre «alienígenas» reflexivos y pacíficos que en realidad trata sobre nosotros. No incluye espectaculares batallas espaciales, ni extraños extraterrestres de formas exóticas: no es *La guerra de las galaxias* ni *Star Trek*. Los estudios lingüísticos dejan las fuerzas militares al margen, y hay un mensaje importante al final de la película que revela por qué vienen las doce naves, y después, también de repente, se marchan y dejan tras de sí únicamente un atisbo de la profunda complejidad de su modelo de realidad.

El valor del mensaje de esta película estriba en que nos hace reflexionar sobre cuánto influye el funcionamiento de la mente y sus conexiones para adoptar ciertas visiones del mundo y adquirir ciertos procesos, como el cómputo.

Los avances en la investigación destinada a proporcionar una descripción rigurosamente lógica de la aritmética han sido relevantes por muchas razones. Han modificado la idea que teníamos del cómputo y de los números y, con ello, dejamos de verlos como herramientas funcionales para contar huevos o monedas. Los números naturales empezaron a existir con independencia de las cosas enumeradas, como un sistema lógico definido tan solo por sus reglas, en oposición a aquellos primeros contadores por unidades o por pares. Cambiando las reglas podíamos inventar nuevos sistemas matemáticos cuyos preceptos no

se correspondieran con nada de lo que hay en el mundo. Lo único importante era que las reglas fueran coherentes y no implicaran la deducción de contradicciones, como que 1 = 2. El concepto de «existencia matemática», que aparece con frecuencia, cobró un significado nuevo. Ya no se refiere a que podamos encontrar ejemplos de ese sistema matemático en la vida real, sino que significa únicamente que posee coherencia interna: este es el significado de existencia matemática.

En la Edad Media se creía que las matemáticas, o más en concreto la geometría euclídea, eran una descripción única del funcionamiento del mundo. La geometría euclídea describía la estructura de los puntos, de las líneas y de las superficies planas. Se consideraba parte de la verdad absoluta sobre las cosas y, como tal, se podía enarbolar ante los escépticos que proclamaban la incapacidad de la mente humana para captar aspectos de la verdad última, como los que solían debatir teólogos y filósofos. Pero el descubrimiento de geometrías no euclídeas en el siglo XIX por parte de Gauss, Bolyai y Riemann, las cuales describen puntos y líneas sobre una superficie curva, como la Tierra, cambió radicalmente aquella concepción [8]. Ahora sabemos que había muchas geometrías posibles, todas ellas definidas con coherencia interna por diferentes conjuntos de axiomas. Todas ellas existen matemáticamente en el sentido en el que acabamos de describir. Más tarde se descubrirían nuevas lógicas y nuevos sistemas aritméticos. Esto generó una especie de relativismo en relación con las matemáticas. De este modo, los grandes volúmenes sobre física matemática con títulos como «La teoría de las ondas», acabaron siendo reemplazados por obras más modernas tituladas, por ejemplo, «Modelos matemáticos de las ondas en movimiento», porque no había una sola teoría matemática de las ondas, sino que podían emplearse distintas herramientas matemáticas o crear otras nuevas para manejar diferentes aspectos del complejo movimiento de las ondas.

A un nivel más profundo, se observa un distanciamiento de aquella idea de la Antigüedad que atribuía a los números significados intrínsecos (el 13 da mala suerte, el 7 es propicio…). Esta fue la base de la tradición de la numerología, una idea tentadora que persiste en algunos círculos actuales, una reliquia de

la mentalidad pitagórica. Así, por ejemplo, se creía que todos los sietes compartían un mismo significado. Las matemáticas modernas que emergieron a partir del trabajo de Peano no atribuyen ningún significado a los números ni a los puntos o líneas de por sí, sino tan solo a las relaciones que mantienen entre ellos. Por tanto, las matemáticas son el estudio de esas relaciones: los patrones que existen en el mundo constituyen el objeto de estudio de las matemáticas aplicadas y la física matemática, mientras que la formación, la coherencia interna y el estudio de los patrones en sí son el contenido principal de las matemáticas puras.

5. Sumas de conjuntos y otras cosas

A set is a set
(*you bet, you bet*)
And nothing could not be a set,
you bet!
That is, my pet
Until you've met
My very special set.

Bruce Reznick [1].

El desarrollo de una concepción mucho más formal de las matemáticas puras —tratando cada una de sus partes como un juego cuyas reglas rigen las deducciones que se pueden realizar a partir de unas premisas iniciales bien definidas (o axiomas)— condujo a la consideración de los números como conjuntos de cosas. Esta perspectiva la introdujo Georg Cantor en 1874 en relación con la aritmética transfinita, o de cantidades infinitas, la cual veremos en el capítulo siete. Desde un punto de vista informal, un conjunto es una colección de cosas que puede incluir objetos matemáticos, como números, objetos no matemáticos, como teteras, o incluso otros conjuntos. Por tanto, el conjunto {1, 2, 4, 7, 9} contiene el conjunto {1, 2, 9} como subconjunto, pero no contiene el conjunto {1, 2, 3}. Con los conjuntos se pueden realizar diversas operaciones.

La unión de dos (o más) conjuntos, A y B, (que se indica mediante el signo $A \cup B$) da como resultado otro conjunto que contiene todos los miembros de A y de B. Por tanto, si A = {a, b, c} y B = {3, 4, 5, a, b}, entonces $A \cup B$ = {a, b, c, 3, 4, 5}.

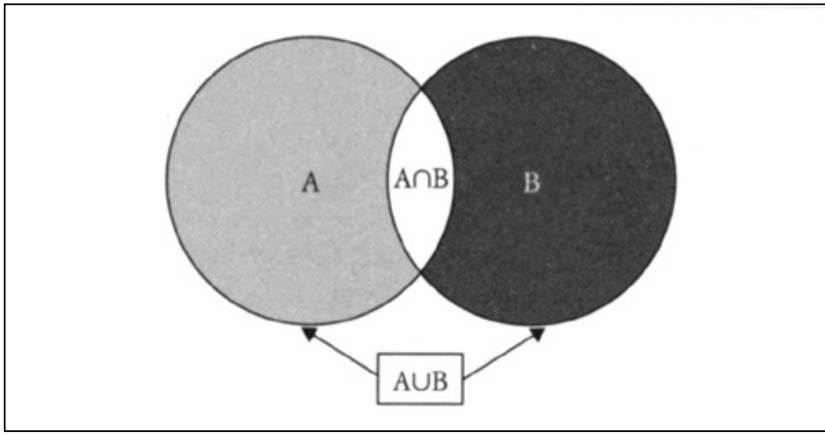


Figura 5.1. La unión y la intersección de los conjuntos A y B.

La intersección de dos conjuntos se indica mediante $A \cap B$ y contiene todos los miembros que son comunes a A y B, de modo que, en el caso de los conjuntos A y B que hemos usado como ejemplo, $A \cap B = \{a, b\}$.

El conjunto vacío [2] no contiene ningún miembro, se representa mediante {} o Ø, y es un subconjunto de cualquier conjunto.

La diferencia entre los conjuntos $A \setminus B$ es el conjunto de todos los miembros de A que no son miembros de B, de modo que, en el caso de los conjuntos que hemos tomado como ejemplo, $A \setminus B = \{c\}$.

El producto de los conjuntos A y B, indicado mediante $A \times B$, es el conjunto de todos los pares que combinan un elemento de A con otro de B, de modo que si $A = \{1, 2\}$ y $B = \{a, b\}$, entonces $A \times B = \{(1, a), (1, b), (2, a), (2, b)\}$.

Otro concepto interesante es el del conjunto potencia, *P(A)*, de un conjunto A, el cual contiene todos los subconjuntos de A. Si $A = \{1, 2\}$, entonces *P(A)* es

$$\{\{\}, \{1\}, \{2\}, \{1, 2\}\},$$

el cual incluye siempre el conjunto vacío, {}, y el conjunto completo, A, que en este caso es {1, 2}. El conjunto potencia de un

conjunto con N elementos contendrá $2N$, por lo que crece muy rápido a medida que aumenta el número de miembros de un conjunto.

Los conjuntos comparten con el ser humano un aspecto curioso. En algunas concentraciones inmensas de datos se esconden muchos subconjuntos de puntos que podrían seleccionarse por tener propiedades compartidas o por formar parte de algún patrón o tendencia en los datos. ¿Hasta qué punto existirían estos subconjuntos, o cualquier conjunto, si ningún observador reconociera y seleccionara lo que tienen en común los elementos que los conforman? Cantor consideraba que los conjuntos existen ya sean reconocidos o no, lo cual sugiere que tendríamos que pensar que ¡el *David* de Miguel Ángel está presente en todo gran bloque de mármol!

En la práctica hay algunas restricciones para las colecciones que pueden considerarse conjuntos. Por ejemplo, el famoso «conjunto de todos los conjuntos» no es un conjunto [3], porque si lo fuera, conduciría a una incongruencia. En esto consiste la célebre paradoja de Russell, la cual plantea si un conjunto puede pertenecer a sí mismo (como ocurriría con el conjunto de todos los conjuntos), y lo descarta para evitar la contradicción. Una versión elemental y mucho más comprensible de esta paradoja la ofrece aquella conocida localidad cuyo único barbero afeita a todos los hombres del lugar que no se afeitan a sí mismos, lo que plantea el interrogante de si el barbero se afeita o no a sí mismo. Un antecedente antiguo de estas dos paradojas lo encontramos en la paradoja del mentiroso, cuyo protagonista declara «Estoy mintiendo», lo que genera la pregunta de si estará mintiendo o no al emitir esa afirmación. Otra variante del mismo argumento considera la frase «Esta afirmación no es verdadera».

La paradoja del mentiroso se relaciona con el filósofo cretense Epiménides de Cnosos, a quien se le atribuye la frase «Todos los cretenses son unos mentirosos» o, más en concreto, «Soy un mentiroso», fechada en torno al año 600 a. C. La encontramos también mencionada en el Nuevo Testamento, en la carta de Pablo de Tarso a Tito. Como es natural, hay que prestar mucha atención al analizar esa afirmación. Por ejemplo, decir en general que todos los cretenses son mentirosos no implica necesariamente que cada uno

de ellos mienta cada vez que habla. En este aspecto, la paradoja pierde todo su poder. Sin embargo, la otra variante, la que dice «Estoy mintiendo», es mucho más sutil y enrevesada.

Estos resultados paradójicos revelan que no todo lo que consideramos una colección de cosas es un conjunto, y entonces se introdujo una noción más amplia y menos rígida: las clases de objetos. Los miembros de una clase (los cuales a su vez podrían ser conjuntos) simplemente poseen una propiedad en común, como tener cuatro patas o formar un conjunto. Reglas axiomáticas precisas inspiradas en la paradoja de Russell y otros razonamientos similares procuran aclarar cuándo una clase puede considerarse o no un conjunto. Así, por ejemplo, prohíben que la clase de todos los conjuntos sea un conjunto.

De este modo, diferenciamos entre conjuntos puros, que son conjuntos a todos los efectos, y clases propias, que no lo son, lo que permite superar la paradoja de Russell.

El conjunto vacío resulta fundamental para definir lo que entendemos por número. En 1923 el polímata John von Neumann reveló cómo hacerlo siguiendo el enfoque con el que Ernst Zermelo abordó este problema en 1908 [4].

Si tenemos en cuenta nuestra definición de conjunto puro, entonces el conjunto puro más simple lo constituye el conjunto vacío, ø, y el siguiente más simple es el que contiene el conjunto vacío como único elemento {ø}. Les sigue el conjunto que contiene tan solo el conjunto vacío más el conjunto que contiene tan solo el conjunto vacío, que es {ø, {ø}}. El siguiente es el conjunto que contiene el conjunto vacío y los dos conjuntos anteriores, que es {ø, {ø}, {ø, {ø}}}, y así sucesivamente. Ahora podemos definir los números naturales usando esta secuencia de conjuntos vacíos anidados que emplea tan solo el conjunto vacío. Por tanto, entre los números naturales y los conjuntos y subconjuntos del conjunto vacío tenemos las siguientes asociaciones:

El número 0 es ø
El número 1 = {0} que es {ø}
El número 2 = {0, 1} que es {ø, {ø}}
El número 3 = {0, 1, 2} que es {ø, {ø, {ø}}}
El número 4 = {0, 1, 2, 3} que es {ø, {ø}, {ø, {ø}}, {ø, {ø},

{ø, {ø}}}},

y así sucesivamente.

Por tanto, ¡ahora hay una manera de crearlo todo a partir de la nada! En contra del antiguo dicho de Parménides de «Nada puede crearse de la nada» (*Ex nihilo nihil fit*), que Shakespeare puso en boca del rey Lear como «Nada saldrá de la nada [5] ». Una definición anterior de Ernst Zermelo, de 1908, persiguió la misma idea recurriendo a una construcción ligeramente distinta de los números naturales a partir del conjunto vacío del siguiente modo: 0 = ø, 1 = {ø}, 2 = {{ø}}, 3 = {{{ø}}}, y así sucesivamente. En las figuras 5.2 y 5.3 se muestran las construcciones de Von Neumann y de Zermelo.

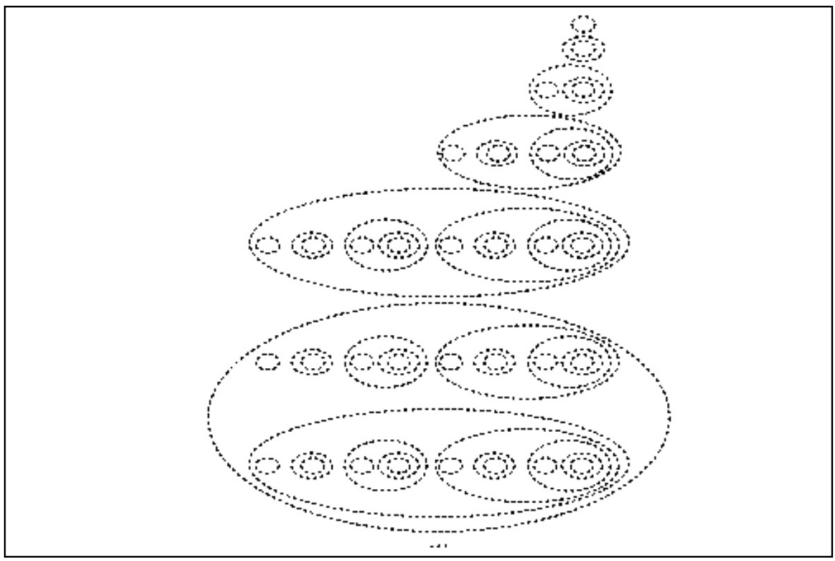


Figura 5.2. Imagen de la construcción de Von Neumann de los números naturales a partir del conjunto vacío.

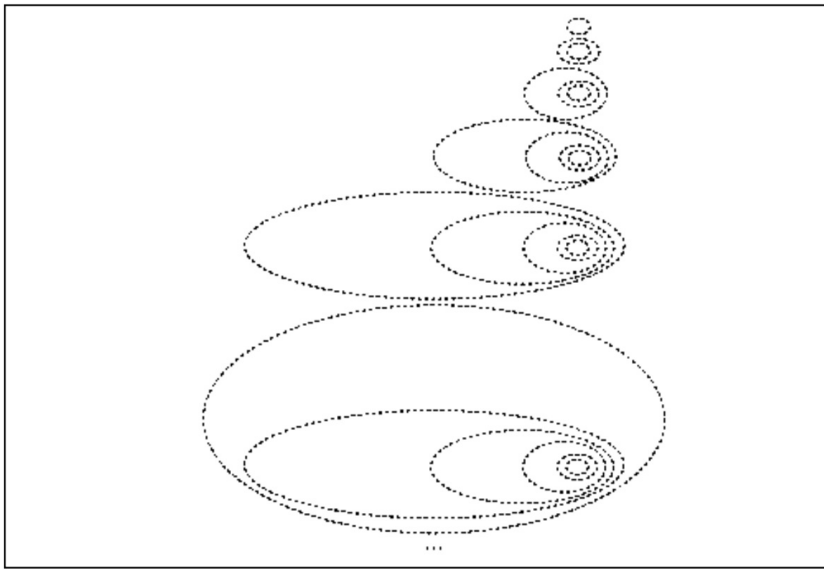


Figura 5.3. Imagen de la construcción de Zermelo de los números naturales a partir del conjunto vacío.

Hay muchas cantidades cuya suma no respeta la fórmula 1 + 1 = 2. Supongamos que actúan sobre usted dos fuerzas de intensidad 1 que mantienen entre sí un ángulo recto. La fuerza resultante que estarán ejerciendo sobre usted no valdrá 2, sino que se corresponderá con la hipotenusa del triángulo que se crea con las dos fuerzas en perpendicular, tal como se muestra en la figura 5.4.

La fuerza resultante, indicada mediante la línea discontinua en diagonal, actúa con un ángulo de 45 grados respecto de las fuerzas vertical y horizontal [6]. Su magnitud M viene dada por el teorema de Pitágoras aplicado al triángulo con la hipotenusa M, donde

$$M^2 = 1^2 + 1^2 = 2$$

De modo que M, la suma de las fuerzas de intensidad 1 y 1, equivale a la raíz cuadrada de 2, que aproximadamente da 1.414.

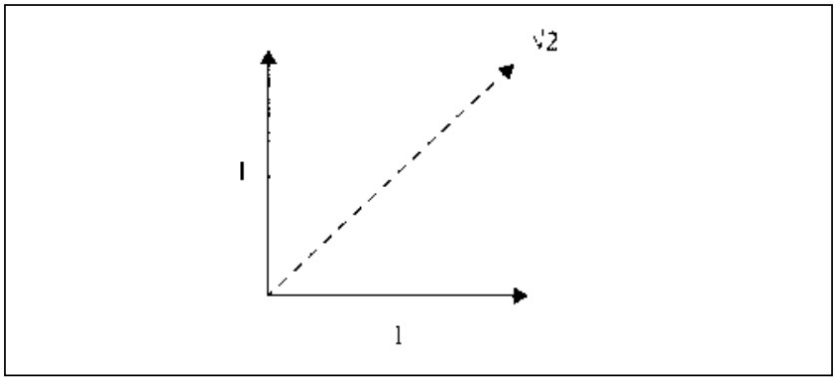


Figura 5.4. La fuerza resultante que se percibe al someterse a dos fuerzas perpendiculares de magnitud 1 es igual a 1.414, no a 2.

6. Demostración de Whitehead y Russell de 1 + 1 = 2

> ¡*Cuánto ha contribuido el trabajo de los lógicos a convertir el inglés en una lengua que permita pensar con claridad y precisión sobre cualquier tema*! Los Principia Mathematica *tal vez supongan una aportación mayor para nuestra lengua que para las matemáticas*.
>
> T. S. Eliot [1].

Bertrand Russell y Alfred North Whitehead fueron miembros del Trinity College, en Cambridge. Whitehead era unos diez años mayor que Russell y había sido profesor suyo. Tras una serie de trabajos más breves, se embarcaron juntos en un gran proyecto, en la redacción de los *Principia Mathematica*, que los ocupó durante diez años. En realidad, no hubo muchos periodos en los que trabajaran de cerca. La mayoría de la obra se gestó entre el otoño de 1906 y el otoño de 1909. Por entonces Whitehead residía en Cambridge y Russell en Oxford [2]. La obra completa se publicó en tres volúmenes en Cambridge University Press entre 1910 y 1913. La composición tipográfica a mano de miles de páginas de lógica matemática tenía un coste prohibitivo (600 libras) y su publicación solo fue posible gracias a una beca de la Real Sociedad de Londres y a las aportaciones personales de Russell y Whitehead. Solo se imprimieron 750 ejemplares. Los volúmenes nunca llegaron a venderse en cantidades significativas, y Russell declaró en cierta ocasión que creía que solo lo habían leído unas seis personas. El precio de venta al público del primer volumen ascendía a 1,25

libras. Acabo de consultar una página en internet especializada en la venta de libros raros (Sophia Rare Books) y he encontrado una flamante primera edición en tapa dura de los tres volúmenes de *Principia* ¡por la que piden 110 000 dólares!

Whitehead y Russell estaban decididos a reducir a la lógica el conjunto de todas las matemáticas y a demostrar su completitud y coherencia usando las reglas deductivas convencionales de la lógica. Quisieron usar la menor cantidad posible de axiomas iniciales sin basar ninguno de ellos en observaciones del mundo natural. En el proceso inventaron numerosos elementos de simbolismo lógico y deductivos. Pero su enfoque, conocido como «logicismo», estaba abocado al fracaso, sobre todo por el trabajo del lógico austriaco Kurt Gödel, con quien volveremos a encontrarnos en el capítulo ocho. En 1931 Gödel demostró que no se puede usar ningún sistema finito, ni lógico ni de ninguna otra clase, para inferir todas las verdades matemáticas si ese sistema es coherente y lo bastante grande como para contener la aritmética.

La obra *Principia Mathematica* supuso un hito intelectual extraordinario. Sin embargo, no pocos aspectos de su presentación son muy ineficaces, con repeticiones y derroteros absolutamente infructuosos. En ningún momento se mencionan los conjuntos, y faltan algunas fórmulas y elementos del entramado lógico que se inventaron más tarde. En consecuencia, los volúmenes nunca se usaron mucho y no llegarían a utilizarse para la enseñanza de la lógica matemática. En la actualidad solo parece interesar a especialistas en historia de las matemáticas y de la lógica. El matemático Godfrey Hardy, compañero de Russell en Trinity College, relató la siguiente anécdota simpática:

> Recuerdo a Bertrand Russell contándome una pesadilla horrible que había tenido. Él estaba en la planta superior de la biblioteca de la universidad hacia el año 2100 d. C. Un ayudante de la biblioteca recorría los anaqueles acarreando un cajón enorme; abría libros, los ojeaba, los devolvía a las estanterías o los echaba al cajón. Al final se detuvo ante tres volúmenes de gran tamaño que Russell reconoció como la última copia existente de *Principia Mathematica*. Tomó uno de los volúmenes, ojeó algunas páginas, por un instante pareció interesado por los curiosos símbolos, cerró el ejemplar, lo sopesó en la mano y vaciló [3] ….

En la actualidad los matemáticos no buscarían una sola lógica ni un conjunto de axiomas único para basar en ellos todas las matemáticas [4]. Saben que pueden crearse distintos sistemas de axiomas en los que encajar diferentes partes de su estudio: la geometría euclídea, la geometría no euclídea, la lógica ternaria, la teoría de grupos y toda clase de álgebras diversas. Desplazaron el imperativo axiomático al campo de la computación o a la escritura de programas coherentes para realizar tareas específicas (rutinas o cálculos). Para instruir a la computadora hay que usar coherencia y unas reglas bien definidas. Algunos de los problemas modernos más conocidos, como el teorema de los cuatro colores [5], los resolvieron por primera vez las computadoras, no personas, debido a la gran cantidad de casos particulares que hay que considerar y comprobar para lidiar con los contraejemplos del teorema.

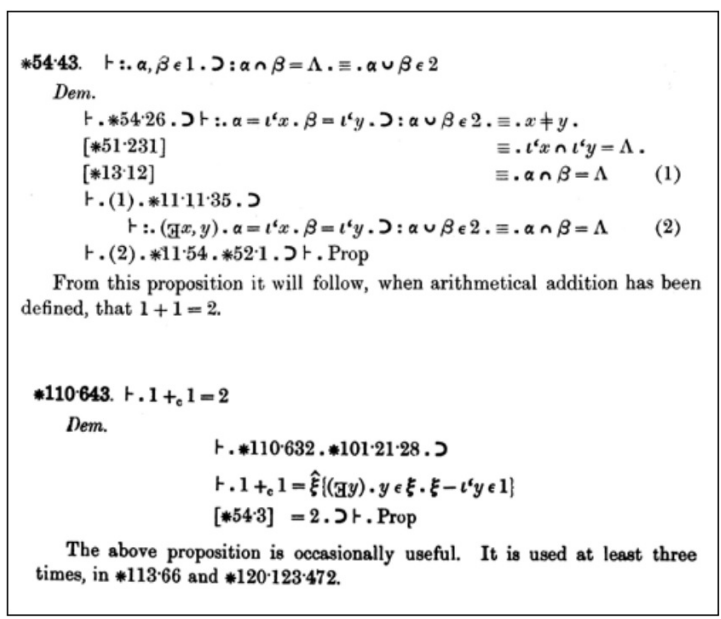
Pero no sucedía lo mismo en la época de Whitehead y de Russell. A mediados del siglo XIX George Boole había formulado la lógica usando álgebra para crear lo que denominamos el álgebra booleana [6], y Frege, Dedekind y Peano habían usado la teoría de conjuntos para definir los números y para formular su adición y multiplicación empleando conceptos de la teoría de conjuntos, como la unión y la intersección de conjuntos, los cuales tenían un alcance más amplio que la simple aritmética [7].

Dentro de este empeño por convertir las matemáticas en operaciones lógicas (en esencia en eso que hoy denominaríamos un programa informático), no es casual que Russell hubiera elaborado previamente una imponente biografía de Leibniz, quien a finales del siglo XVII había intentado crear una máquina que diera respuesta a cualquier pregunta que se le formulara, ya fuera matemática, teológica o política.

Una de las enseñanzas de aquellas iniciativas fue que los sistemas matemáticos se pueden definir de cualquier manera que se desee siempre y cuando mantengan la coherencia. No es necesario que la formulación elegida sea aplicable al mundo físico ni que respondan a una motivación. Son como juegos inventados libremente. Veamos un ejemplo creado por nosotros. Supongamos que tenemos cuatro cantidades llamadas John, Pino, Alessia y Jo, y definimos las reglas para multiplicarlas y sumarlas entre sí en las tablas 6.1 y 6.2.

La razón por la que exponemos aquí el complejo trabajo de lógica de Russell y de Whitehead estriba en que a partir de axiomas simples demuestra que 1 + 1 = 2. La demostración aparece tras cientos de páginas repartida en los dos primeros volúmenes. Por supuesto, por el camino desarrollan y demuestran muchas otras cosas.

La figura 6.1 permite hacerse una idea del tono en que están escritos los *Principia*, y reproduce la parte del texto centrada en la demostración de 1 + 1 = 2, la cual intentaremos hacer inteligible a continuación para legos en la materia. En su lógica, Russell y Whitehead emplean un conjunto de símbolos y notaciones completamente distinto del que usamos hoy día. Pero si traducimos las ideas clave a los términos actuales, podemos ver el funcionamiento de su lógica.



***54·43.** $\vdash :. \alpha, \beta \in 1 . \supset : \alpha \cap \beta = \Lambda . \equiv . \alpha \cup \beta \in 2$

Dem.

$\vdash . *54·26 . \supset \vdash :. \alpha = \iota`x . \beta = \iota`y . \supset : \alpha \cup \beta \in 2 . \equiv . x \neq y .$

[*51·231] $\equiv . \iota`x \cap \iota`y = \Lambda .$

[*13·12] $\equiv . \alpha \cap \beta = \Lambda$ (1)

$\vdash . (1) . *11·11·35 . \supset$

$\vdash :. (\exists x, y) . \alpha = \iota`x . \beta = \iota`y . \supset : \alpha \cup \beta \in 2 . \equiv . \alpha \cap \beta = \Lambda$ (2)

$\vdash . (2) . *11·54 . *52·1 . \supset \vdash .$ Prop

From this proposition it will follow, when arithmetical addition has been defined, that $1 + 1 = 2$.

***110·643.** $\vdash . 1 +_c 1 = 2$

Dem.

$\vdash . *110·632 . *101·21·28 . \supset$

$\vdash . 1 +_c 1 = \hat{\xi}\{(\exists y) . y \in \xi . \xi - \iota`y \in 1\}$

[*54·3] $= 2 . \supset \vdash .$ Prop

The above proposition is occasionally useful. It is used at least three times, in *113·66 and *120·123·472.

Figura 6.1. Extracto de los *Principia Mathematica* donde se demuestra que 1 + 1 = 2.

Ellos usan puntos en lugar de paréntesis para indicar el modo en que se anidan conjuntos dentro de conjuntos (esta notación la tomaron de Peano, quien se la enseñó a Russell cuando se encontraron durante un congreso en París). Es más simple cuando hay muchas anidaciones, pero no nos resulta familiar. De modo que, en lugar de

$$((1 + 3) \times 4) + 6,$$

en los *Principia* encontramos

$$1 + 3 . \times 4 : 6.$$

El símbolo : es en realidad un punto doble que indica que toda la fórmula está aislada, de modo que el 6 se suma a todo lo anterior, lo que da como resultado 22. En fórmulas más complejas podemos encontrar tres o cuatro (o más) puntos indicados mediante :., .: o ::, y así sucesivamente.

Lo que pretende ilustrar la primera línea de la figura 6.1, en la proposición 54.43, es que si cada uno de los conjuntos α y β cuenta exactamente con un elemento, entonces están desunidos si y solo si su unión tiene exactamente dos elementos. La demostración, a partir de la abreviatura *Dem.* (de «demostración») en la figura, indica que de la proposición 54.26 se deduce que si $\alpha = \{x\}$ y $\beta = \{y\}$, entonces $\alpha \cup \beta$ tiene 2 elementos si y solo si x es diferente de y. Entonces, de la proposición 51.231 se deduce que «x es diferente de y» es verdadero si y solo si los conjuntos {x} e {y} están desunidos y, recurriendo a la proposición 13.12, $\alpha \cap \beta = \Lambda$ (que es el conjunto vacío Ø).

Llegados a este punto, Whitehead y Russell concluyen con el resultado etiquetado como (1), que si $\alpha = \{x\}$ y $\beta = \{y\}$, entonces $\alpha \cup \beta$ tiene 2 elementos si y solo si $\alpha \cap \beta = \Lambda$. Este teorema aparece etiquetado como (2) y constituye lo que se pretendía demostrar en la proposición 54.43. En esencia evidencia que $1 + 1 = 2$, porque ya vimos que si α y β contienen un solo elemento cada uno, entonces su unión (suma) contiene 2 elementos si son disjuntos. Este último paso se da en el segundo extracto, proposición 110.643, de la figura 6.1. El cálculo de $1 + 1$ requiere

encontrar dos ejemplos disjuntos de 1 y considerar su unión; el teorema 54.43 establece que la unión debe ser un elemento de 2, con independencia de los ejemplos escogidos, de modo que al final $1 + 1 = 2$. Los autores señalan con modestia que el resultado $1 + 1 = 2$ ¡es «útil en ocasiones»!

Se trata de una materia compleja, y no continuaremos con ella aquí. De hecho, Russell notó que su mente se resintió por la densidad de este trabajo. Ni Whitehead ni Russell volvieron a investigar nunca más en detalle este tema sobre los fundamentos de las matemáticas.

7. Aritmética transfinita

El infinito convierte lo posible en lo inevitable.

Norman Cousins [1].

Una de las reflexiones más profundas sobre el infinito durante la Antigüedad fue la de Aristóteles [2]. Él diferenciaba entre dos tipos de infinito: el «infinito en potencia», que es como los números naturales 1, 2, 3, 4… y así eternamente, o la secuencia infinita hacia atrás de los números negativos… –6, –5, –4, –3, –2, –1, cuyo punto de partida es inalcanzable. La secuencia es interminable (basta con añadir un número más cuando creas haber llegado al final): es infinita. En este caso se dice que «tiende al» infinito porque nunca llega hasta él. Un universo de espacio ilimitado también es así. Hay un volumen infinito que nunca llegaremos a alcanzar. Todos estos son ejemplos de infinitos en potencia, y Aristóteles estaba bastante feliz con su existencia. Lo que lo tenía bastante descontento eran los «infinitos en acto». Estos se corresponden con infinitos dentro de cantidades locales medibles y observables, como la temperatura, el brillo, la densidad de la materia, la fuerza o la velocidad. Tenían la capacidad de influir en nosotros y destrozar otros aspectos de la cosmovisión aristotélica. Aristóteles proscribió la existencia de un vacío local perfecto, puesto que no opondría ninguna resistencia al movimiento, lo que permitiría a los objetos alcanzar una velocidad infinita. Pero sí admitió el infinito en potencia de las secuencias de tiempo pasadas y futuras, porque no quería que el tiempo tuviera principio ni fin. Creía que la materia se puede dividir hasta el infinito. Por ejemplo, entre dos puntos distintos siempre hay uno intermedio, de la misma manera que entre dos números reales, o incluso entre dos fracciones

x < y, siempre se puede insertar la semisuma (x + y) / 2. Pero, además de la densidad, Aristóteles también creía en el carácter continuo de la línea recta. En esto se oponía a los atomistas, que defendían que hay unidades mínimas de materia (los átomos) a partir de las cuales se creó todo lo demás.

Cuando la filosofía aristotélica se fundió con la teología de la Iglesia católica en la Edad Media, hubo dos puntos de fricción cruciales entre el infinito y esa doctrina religiosa. El primero y más evidente era que solo Dios es infinito. Cualquier aceptación de otros infinitos (matemáticos o de otra índole) suponía un desafío a esa doctrina y se consideraba herético. Descartes, en el pasaje siguiente, manifiesta ese prejuicio en contra del estudio del infinito:

> reservar el término *infinito* únicamente para Dios, porque únicamente en el caso de Dios no solo no se admite ningún límite en ningún sentido, sino que nuestro entendimiento nos dice positivamente que no lo hay [3].

En cambio, otro pensador perspicaz francés como fue Blaise Pascal puso gran énfasis en la doble infinidad que encontramos en la naturaleza: lo infinitamente grande, pero también la relevancia, a menudo inadvertida, de lo que posee una pequeñez infinitesimal. Esto último es mucho más espectacular porque existe en el interior de cualquier fragmento de materia, incluso en la palma de la mano [4].

El otro desacuerdo se centró en discernir si los infinitos en potencia, como la secuencia interminable de los números, podían cuestionar la omnipotencia de Dios. ¿Podía Dios determinar el fin de una serie y preservar así la naturaleza divina del infinito? San Agustín pensaba que sí y defendía que la sabiduría de Dios era capaz de hacer frente al desafío del infinito:

> En cuanto a la afirmación de que ni tan siquiera la omnisciencia de Dios es capaz de comprender lo infinito, lo único que le falta a quienes defienden esto es hundirse en las profundidades de su impiedad atreviéndose a decir que Dios no conoce todos los números. En efecto es cierto que los números son infinitos… ¿Significa esto que Dios no conoce todos los números porque son infinitos? ¿Acaso la sabiduría de Dios llega hasta cierta cantidad y se acaba ahí? Nadie podría estar lo

> bastante loco para afirmar esto… De modo que no dudemos jamás que conoce todos y cada uno de los números… de lo que se deriva por necesidad que toda infinitud se vuelve de algún modo finita para Dios, porque no puede quedar fuera del alcance de su omnisciencia [5].

Este tipo de problemas alcanzó su punto álgido en los primeros debates medievales sobre si las acciones de Dios están limitadas por las leyes de la naturaleza y de la lógica. La conclusión fue que Dios podía hacer todo lo que fuera lógicamente posible (de modo que no podía hacer que $1 + 1 = 3$ en la aritmética convencional) y podía saber todo lo cognoscible. Por tanto, el resultado de tirar un dado al aire podía considerarse incognoscible en principio antes de efectuar el lanzamiento (aunque tal vez no, si se cuenta con una información muy detallada sobre la dinámica de ese lanzamiento [6]).

Aun así, la aversión a permitir el infinito dentro de las matemáticas perduró hasta mucho tiempo después de la Edad Media. Grandes matemáticos del siglo XIX, como Karl Friedrich Gauss, seguían insistiendo en que los únicos infinitos existentes eran los que Aristóteles consideraba potenciales. No podía haber infinitos físicos ni manipulaciones matemáticas que incluyeran el infinito, porque podían dar lugar a conclusiones extrañas como

$$1 + \infty = \infty = 2 + \infty, \text{ de manera que } 1 = 2.$$

Varios matemáticos influyentes temían que la inclusión del infinito en las matemáticas pudiera conducir al colapso de toda su estructura lógica porque si se permite una falacia en un sistema lógico, entonces cualquier afirmación puede demostrarse como cierta. Esto propició el desarrollo de un enfoque finitista de las matemáticas que solo admitía como «verdaderos» aquellos teoremas demostrados mediante una cantidad finita de pasos a partir de los axiomas iniciales de la aritmética. Es lo que se denominó a veces constructivismo, del cual hablaremos en el capítulo diez. Esta definición de la verdad matemática da como resultado unas matemáticas más reducidas que las que utilizamos en la actualidad, porque excluyen las demostraciones no constructivas. Entre estas últimas se cuentan aquellas demostraciones en las que se presupone que algo es verdadero para después inferir una contradicción lógica a partir de ese aserto y concluir, por tanto, que el supuesto de

partida no puede ser cierto. O bien se demuestra que algo tiene que existir sin identificarlo de manera específica ni exponer cómo se construye. Esta filosofía abrió una gran brecha en las matemáticas porque los mayores defensores del constructivismo ocupaban posiciones influyentes, como la dirección editorial de revistas y presidencias de academias. Equipos editoriales completos dimitieron como protesta contra este intento de interferir en el significado de la verdad matemática y de excluir la publicación de matemáticas no constructivistas.

Una de las víctimas de esta disputa fue el matemático alemán cuyo trabajo, ahora clásico, convirtió en riguroso y fascinante el estudio de los infinitos. Georg Cantor cayó en tal depresión ante la nula acogida que tuvo su obra entre los guardianes de las publicaciones matemáticas, que entró en crisis y dejó las matemáticas por un tiempo para trabajar en los orígenes del cómputo (véase el capítulo dos del presente libro), potenciar sus magníficas dotes como dibujante y artista, y hasta para ahondar en la controvertida cuestión de si Shakespeare fue en verdad el único autor de las obras que se le atribuyen.

Cantor reveló cómo definir correctamente un conjunto infinito y después mostró que hay una jerarquía interminable de infinitos cada vez más grandes, señalando con precisión qué significa «más grande» en este contexto. Desveló cómo es una aritmética de cantidades infinitas, la denominada aritmética transfinita. Como veremos, esta aritmética estaba sujeta a unas reglas diferentes a las de la aritmética de los números finitos que definió Peano y que usamos en la vida cotidiana. Para combinar infinitos (para lo que usaremos de manera informal el símbolo ∞ por ahora), la nueva aritmética arrojaba los siguientes resultados:

$$\infty + 1 = \infty,$$
$$\infty + \infty = \infty,$$
$$\infty - 1 = \infty,$$
$$\infty \times \infty = \infty.$$

Cantor se inspiró en las ideas de Galileo, quien había detectado un rasgo extraño en la práctica del cómputo [7]. Galileo había señalado que la relación de los números naturales 1, 2, 3, 4, 5… es infinita, y también lo es la lista de sus cuadrados 1, 4, 9, 16, 25…

diciendo:

> […] ya que hasta cien hay diez cuadrados, que es tanto como decir que solo la décima parte son cuadrados, y en diez mil solo la centésima parte son cuadrados, mientras que en un millón la cifra ha descendido a la milésima parte [8].

La lista de los cuadrados será sin duda más pequeña que la de los números naturales porque solo es un subconjunto de la primera lista de todos los números. Todo esto parece tener lógica. Sin embargo, Galileo dice que si relacionamos los números de la primera lista con los de la segunda del siguiente modo

$$1 \leftrightarrow 1$$
$$2 \leftrightarrow 4$$
$$3 \leftrightarrow 9$$
$$4 \leftrightarrow 16$$
$$5 \leftrightarrow 25$$
$$\dots$$

y así sucesivamente, asociando cada entero con su cuadrado, entonces a cada miembro de la lista de los números naturales (en la columna de la izquierda) le corresponderá uno y solo uno de los miembros de la lista de sus cuadrados (en la columna de la derecha). De modo que ¡ambas columnas tienen que contener la misma cantidad de elementos! Galileo llegó a la conclusión de que el concepto de infinito crea paradojas y que es mejor dejarlo aparte, porque «no podemos hablar de cantidades infinitas más grandes, más pequeñas o iguales a otras». Existen otros ejemplos simples para probar: todos los números naturales pares (2, 4, 6, 8…) o todos los impares (1, 3, 5, 7…) se pueden utilizar en lugar de los cuadrados que propuso Galileo.

Hasta el formidable intelecto de Isaac Newton fracasó al enfrentarse a los misterios del infinito. Él creía, en efecto, que en la paradoja de Galileo había menos cuadrados que números, aduciendo (equivocado, tal como había evidenciado Galileo y tal como revelaría con toda claridad Cantor) que

> aunque hubiese un número infinito de pequeñas partes finitas en una

> pulgada, aún hay doce veces ese número de partes en un pie. Esto es, el infinito número de esas partes en un pie no es igual al de la pulgada, sino que es doce veces más grande que el infinito número de partes que hay en una pulgada [9].

En 1873 Cantor reparó en que el tipo de paradoja que dejó tan perplejo a Galileo y que este último contempló como un problema insalvable para intentar manejarse con cualquier infinidad es justamente todo lo contrario. Los ejemplos de Galileo y Newton comparten una característica sencilla. Representan dos colecciones (que podemos llamar conjuntos) que dan lugar a una paradoja aparente porque uno de ellos contiene al otro como subconjunto (por ejemplo, todos los números naturales contienen sus cuadrados como subconjunto). Esto se tomó como base para definir que un conjunto es infinito si se puede poner en correspondencia de 1 a 1 con alguno de sus subconjuntos, tal como sucede en el ejemplo de Galileo entre los números naturales y sus cuadrados respectivos. Fue Dedekind, y no Cantor, quien propuso esta caracterización. Cantor prefirió una diferente, y denominó infinito a aquel conjunto que no se puede poner en correspondencia de 1 a 1 con {1, 2..., n} para ningún entero positivo n. Parecen dos formas distintas de decir lo mismo, pero su equivalencia no es nada fácil de demostrar.

Cantor se dio cuenta de que esta correspondencia de uno a uno de los conjuntos con los números naturales equivale sencillamente a contar los miembros de esos conjuntos, y denominó cardinalidad a la cantidad, seguramente infinita, de esos miembros. En primer lugar, definió como numerables aquellos conjuntos infinitos que pueden ponerse en correspondencia de uno a uno con los números naturales y denotó esa cardinalidad mediante la letra alef del alfabeto hebreo, א, seguida de un cero, 0, en subíndice. Pero este solo fue el primer paso de un viaje intelectual espectacular que acabaría desvelando que existe una torre infinita de infinitos cada vez más y más grandes (de mayor cardinalidad). Por un lado, Cantor demostró que el conjunto de todas las fracciones (o números racionales, tal como las llamaban los griegos) es un infinito numerable. El truco consiste en saber cómo contar estos números. Recordemos que estos se representan de manera única con la forma p / q donde p y q son números enteros primos entre sí, es decir, sin

más divisores comunes que 1 y –1, y que q es positivo. En caso contrario, la fracción se simplifica, como sucede con 2/4, que equivale a 1/2, o con 4/2, que es 2 o, si se prefiere, 2/1. El método de Cantor para contarlas, llamado diagonal, reordena las fracciones simplificadas p / q (o al menos las que son positivas) de acuerdo con la suma $p + q$: así que se parte de fracciones donde $p + q = 2$ (es decir, tan solo 1/1), se sigue con $p + q = 3$ (o sea, con 1/2 y 2/1); luego con $p + q = 4$ (el caso de 1/3 y 3/1, porque 2/2 se simplifica y ya ha aparecido como 1/1); después viene $p + q = 5$ (1/4, 2/3, 3/2, 4/1), etc. Solo hay un número finito de fracciones que comparten el mismo resultado al sumar $p + q$. Si esta suma da el mismo resultado, se da prioridad a p, es decir, se sigue el orden convencional de las fracciones. De este modo se obtiene la siguiente lista:

1/1, 1/2, 2/1, 1/3, 3/1, 1/4, 2/3, 3/2, 4/1…

que ordena las fracciones positivas de la misma manera que 1, 2, 3, 4…, de modo que mantienen una correspondencia de 1 a 1 con los números naturales. Ambos conjuntos son numerables, con cardinalidad alef sub cero. Adaptaciones sencillas amplían el resultado a todas las fracciones, incluso a las que valen cero o son negativas. Sin embargo, ya en 1874 Cantor había demostrado que otro conjunto de números relevantes, el de los números reales, tiene una cardinalidad diferente a la de los numerables. Su teorema tuvo unas consecuencias históricas, tanto más cuanto que con el paso de los años Cantor lo mejoró y descubrió, como ya se ha dicho, no uno, sino una infinidad de infinitos no numerables: abismos tan vertiginosos que dificultan una clasificación.

Cantor se puso a analizar lo que en matemáticas se conoce como números reales, que incluyen aquellos que no pueden expresarse como fracciones del estilo p/q. Esto abarca los números cuyo desarrollo decimal no tiene fin, como, por ejemplo, ciertas fracciones racionales (como $1/3 = 0.333\,333\,33…$), pero también números famosos como $\pi = 3.141\,592\,653\,5…$ o el número $e = 2.718\,281\,828\,4…$

A continuación, Cantor procedió a demostrar que el conjunto infinito de todos los números decimales que no terminan nunca no

puede tener una cardinalidad numerable, no son contables, se trata de una cantidad mayor que $\aleph_0$. Lo que hizo fue comprobar que toda la recta real tiene la misma cardinalidad que un segmento de ella, aunque sea pequeño, como el intervalo entre 0 y 1 (ambos excluidos). Quienes estén familiarizados con la trigonometría y con la gráfica de la función tangente podrán comprobarlo con facilidad. Así que esto basta para excluir que los números reales entre 0 y 1 se puedan ordenar con una correspondencia de uno a uno con los números naturales. Pero supongamos pese a todo que sí existiera tal correspondencia y que, por lo tanto, los números reales entre 0 y 1 se pudieran listar en paralelo con 0, 1, 2, 3… Recordemos que los números reales tienen una representación decimal, como, por ejemplo, 0.236 109 84…, y que esta representación es única excepto en casos como 0.240 000… = 0.239 999…: no entraremos a explicar el porqué, pero baste decir que la regla subyacente es que 0.999… es igual a 1. En cualquier caso, los números 0 y 9 pueden causar cierta confusión y son un signo de peligro.

Admitamos para simplificar que la lista resultante de números entre 0 y 1 es la siguiente. En realidad, no importa cómo se componga, porque el argumento se puede generalizar con facilidad para cualquier criterio que se adopte. Prestemos atención a los dígitos subrayados, que se desplazan hacia la derecha al bajar en la lista.

0 → 0.235 667 89…
1 → 0.575 603 66…
2 → 0.463 775 21…
3 → 0.846 213 40…
4 → 0.562 106 28…
5 → 0.466 732 30…

y así sucesivamente. La secuencia de las cifras subrayadas es 2, 7, 3, 2, 0, 2, etc. Consideremos ahora un número cualquiera entre 0 y 1 que tenga

- la primera cifra decimal diferente de 2, y también, para evitar confusiones, de 0 y 9,
- el segundo dígito distinto de 7 y también, por la misma razón

que antes, de 0 y 9,
- el tercero que no sea 3, 0, 9,
- el cuarto que no sea 2, 0, 9,
- el quinto distinto de 0 y 9,
- el sexto diferente de 2, 0, 9…

Obsérvese que cada vez se excluyen como máximo 3 cifras de las 10 disponibles y, por lo tanto, quedan al menos 7 disponibles. Pero un número construido de este modo no puede aparecer en ningún lugar de la lista porque, por el modo en que se ha confeccionado, difiere de cualquier número que ya estuviera en ella en al menos un dígito.

Por tanto, se concluye que los números reales no pueden contarse de un modo sistemático. Constituyen un infinito no numerable. Este tipo de infinito suele denominarse el continuo, y su cardinalidad se denota, por razones que veremos más adelante, mediante un dos elevado a alef sub cero: $2^{\aleph_0}$.

Con este golpe maestro, Cantor evidenció que existen infinitos diferentes que no se pueden poner en correspondencia de uno a uno entre sí. Este es uno de los mayores descubrimientos dentro de las matemáticas. Hay motivos para pensar que el continuo es el infinito más grande que se puede encontrar en las aplicaciones científicas. Su cardinalidad es, de hecho, la misma que la del conjunto de los números reales y, tal como podría demostrarse, también que la del conjunto de los números complejos.

Los infinitos superiores pertenecen al ámbito de la lógica matemática, ya que Cantor fue capaz de demostrar que, de acuerdo con las premisas axiomáticas adecuadas, hay una torre infinita de infinitos distintos. Decimos «más grandes» justamente para enfatizar que el continuo no es el fin de la historia.

Para crear esos infinitos nuevos, Cantor hace referencia al conjunto potencia de un conjunto, que es el conjunto de todos sus subconjuntos y que siempre conduce a un infinito más grande que el conjunto en sí. En el ámbito finito, si un conjunto tiene N elementos, su conjunto potencia tendrá 2^N elementos. Por ejemplo, el conjunto {1, 2, 3} tiene como conjunto potencia

$$\{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$$

que tiene $2^3 = 8$ miembros.

Pero esta misma situación se manifiesta en el infinito. Tal como también demostró Cantor, ningún conjunto mantiene una correspondencia de 1 a 1 con el conjunto formado por sus partes. Por ejemplo, el conjunto de números naturales es numerable, pero el de sus subconjuntos tiene la cardinalidad del continuo, la misma que el de los números reales. Entre otras cosas, este descubrimiento justifica la forma en que se denota la cardinalidad del continuo, como $2^{\aleph_0}$, o sea, 2 elevado a la cardinalidad de los números naturales, en analogía con lo que sucede en el caso finito con los conjuntos potencia.

La construcción puede repetirse entonces hasta el infinito, pasando de cada nuevo conjunto potencia así obtenido a su conjunto potencia particular, es decir, al conjunto de sus subconjuntos. La cardinalidad así obtenida se revela cada vez más grande.

Sin embargo, nada excluye que se pueda generar otra cardinalidad prescindiendo del paso al conjunto potencia. Para intentar ordenarlas todas, hay que aclarar en primer lugar el concepto de conjunto, tal vez procediendo a axiomatizarlo.

Pero, tras estas aclaraciones de partida, Cantor mostró que después de lo numerable viene la torre ya mencionada, es decir, una escala ascendente de infinitos, cada uno de ellos superior a los anteriores y ninguno con una correspondencia de 1 a 1 con los demás. Después de $\aleph_0$ vienen $\aleph_1$ (el primero que le sigue), $\aleph_2$, y así sucesivamente [10].

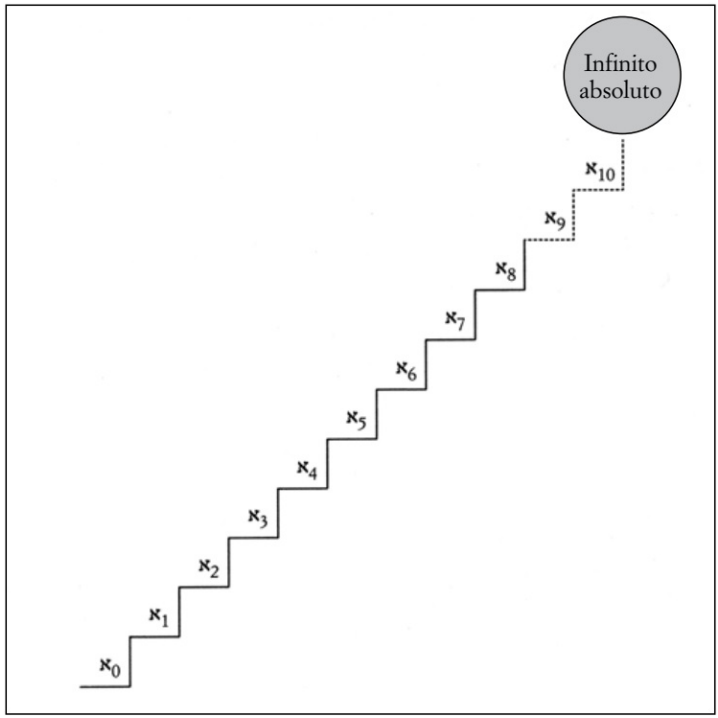


Figura 7.1. La torre ascendente de infinitos cada vez más grandes de Cantor.

De modo que la torre de infinitos de Cantor es interminable, y la asíntota que se forma después de una serie infinita de pasos se denomina infinito absoluto, el cual se representa mediante la letra omega mayúscula, Ω. Sin embargo, este infinito absoluto, u Ω, ya no pertenece a la esfera de la ciencia, sino que, según Cantor, se corresponde de alguna manera con la infinitud trascendente de Dios. Cantor, que era un luterano devoto, puso especial cuidado en diferenciar el infinito supremo de la fe de los dos infinitos permitidos en las matemáticas: el infinito impropio o potencial, que ya había admitido Aristóteles, y su nuevo infinito de los conjuntos y de la cardinalidad, que él llamó infinito propio:

> El infinito en acto se distinguía por tres relaciones: en primer lugar, en tanto que se realiza en la más alta perfección, en el ser extraterrenal completamente independiente, en Dios, donde lo llamo infinito absoluto

> o infinito sin más; en segundo lugar, en tanto que está representado en el mundo dependiente de las criaturas; en tercer lugar, en tanto que puede concebirse en abstracto con el pensamiento como magnitud matemática, como número o como una clase de orden [11].

Cantor parece concebir el infinito absoluto igual que hizo san Anselmo en su célebre argumento ontológico sobre la existencia de Dios invocando la idea de que es más grande y está por encima del resto de cosas que se pueden concebir [12].

Por desgracia, las grandes ideas de Cantor no fueron bien recibidas por los guardianes conservadores de las publicaciones matemáticas más destacadas. Los finitistas las rechazaron de plano y acusaron a Cantor de ser un «corruptor de jóvenes» y un «charlatán científico [13]». Leopold Kronecker, uno de los más influyentes, respondió escribiendo que

> Las definiciones tienen que contener la forma de llegar a una decisión en una cantidad finita de pasos, y las pruebas de existencia tienen que ejecutarse de manera que la cantidad en cuestión se pueda calcular con el grado de precisión necesario [14].

Incluso Carl Gauss, el matemático más grande de aquel tiempo (y tal vez de todos los tiempos), escribió lo siguiente a un amigo en 1831 en relación con el manejo de infinitos en acto:

> Protesto contra el uso que se hace de una cantidad infinita como una cantidad completa, lo que en matemáticas no está permitido. El infinito no es más que una forma de hablar, pero el concepto subyacente real es el de límite, de modo que ciertas fracciones se acercan cada vez más a un valor mientras que a otras les está permitido crecer sin restricción.

Antes de Cantor, la mayoría de matemáticos, como Gauss, consideraban significativos tan solo los infinitos en potencia, pero no los infinitos en acto. La manipulación que realizaba Cantor de los infinitos en acto en su obra de 1883 fue motivo de anatema. Cantor no consiguió publicar su trabajo, y su convencimiento de que Kronecker (quien afirmó que «Dios creó los números enteros, todo lo demás es obra del hombre») lo bloqueaba a cada paso, lo condujo a una crisis nerviosa en 1884 que lo obligó a pasar varios

periodos confinado en una clínica de Halle, ciudad en cuya universidad enseñaba Cantor. Aquello lo apartó de las matemáticas, lo centró en el estudio de la historia del origen de los sistemas de cómputo y lo llevó a interaccionar más con filósofos y teólogos sobre la naturaleza del infinito.

Aunque la mayoría de especialistas en matemáticas sentía poca simpatía por el trabajo de Cantor, la Iglesia católica, sobre todo a través de la influyente figura del sacerdote, filósofo y teólogo alemán Constantin Gutberlet, lo recibió como un regalo del cielo. Sin embargo, había otros teólogos que acusaban a Cantor de crear una nueva forma de panteísmo. Los diferentes tipos de infinito permitían que el infinito de Dios se situara en la cúspide de la torre de Cantor como infinito absoluto y legitimaban el estudio del resto de infinitos inferiores dentro de las matemáticas y de cualquier otro ámbito, sin desafiar la unicidad del infinito divino. Además, esto evidenciaba que la mente humana era capaz de comprender y estudiar esos aspectos del infinito superlativo de lo Divino. Cantor llegó a afirmar incluso que todos los pensamientos inmutables de Dios tenían que ser un infinito completo, lo que ofrecía una prueba de que existen los infinitos superiores de Cantor.

El infinito matemático no era más que un tipo de infinito. Podría haber infinitos físicos (en el momento de la Gran Explosión o en el centro de agujeros negros) o un tipo más general de infinito trascendental que se parezca a Dios o al universo. Diversos matemáticos y filósofos creyeron a menudo en la existencia de alguno de estos tipos de infinito, pero no en la de otros. La tabla 7.1 compendia las creencias de algunos pensadores eminentes sobre la realidad de las tres variedades de infinito [15]:

Este desarrollo extraordinario de las matemáticas condujo a Cantor a la creación de la aritmética transfinita, que trata los infinitos como objetos reales, no como meras potencialidades, con lo que estableció las nuevas y extrañas reglas de la aritmética transfinita que comentamos en los párrafos iniciales de este capítulo. Estas reglas, que rigen los infinitos entre los distintos alef (a partir de $\aleph_0$), especifican y aclaran las reglas que habíamos expuesto con anterioridad de un modo algo tosco mediante el

símbolo ∞.

Así, por ejemplo, se puede ver que $\aleph_0 + 1$, igual que $\aleph_0 + 2$, son lo mismo que $\aleph_0$. De hecho, añadir uno o dos elementos a un conjunto numerable no altera su cardinalidad. Sin embargo, esto no conduce a que pueda demostrarse que $1 = 2 = 0$.

Cantor creía firmemente que estos números transfinitos y los infinitos asociados fueron descubiertos (que existen de verdad), no inventados. Pensaba que le habían sido revelados por intermediación divina. Su trabajo acabó formando parte de las matemáticas convencionales gracias al apoyo del mayor matemático de finales del siglo XIX y principios del XX, David Hilbert. Este se oponía a la doctrina de finitistas como Kronecker, porque consideraba que sus restricciones reducían el alcance de las matemáticas. En 1926 escribió enérgico: «Nadie será capaz de expulsarnos del paraíso que Cantor creó para nosotros [16] ».

Por desgracia, Cantor falleció en 1918 sin llegar a ver la rehabilitación ni la inmensa influencia de su revolucionario trabajo.

8. La incompletitud de Gödel

No es lo mismo una cosa que hablar de una cosa.

Kurt Gödel

Es larga la historia de la consideración científica y filosófica de los imposibles físicos [1]. La cosmovisión aristotélica prohibía que pudieran formarse u observarse infinitos físicos o vacíos físicos locales [2]. Durante la Edad Media hubo físicos que idearon ingeniosos experimentos mentales para intentar imaginar cómo «engañar» a la naturaleza y permitir la aparición de un vacío instantáneo, y explicar después de qué manera se evitaba que ocurriera esa posibilidad mediante procesos naturales o, si eso fallaba, invocando a un Censor Cósmico que impidiera su aparición [3], aunque fuera de manera momentánea.

En la actualidad los especialistas en física de la gravitación que estudian los agujeros negros y la cosmología tratan de demostrar que la teoría de la relatividad general de Einstein no permite la aparición de infinitos físicos, salvo al principio y al final del universo, a no ser que se encuentren ocultos tras el horizonte de los agujeros negros, donde no pueden influir en el universo exterior.

La química vivió su propio debate alquímico sobre la posibilidad o imposibilidad de fabricar oro a partir de metales menos nobles, y dentro de la ingeniería perduró el gusto por buscar máquinas de movimiento perpetuo hasta que la interpretación sistemática de las consecuencias de las leyes de la termodinámica a lo largo del siglo XIX lo erradicó por completo. Ejemplos sutiles, como el del demonio de Maxwell, permanecieron sin solución hasta quedar exorcizados por la moderna termodinámica computacional en 1961 [4].

Los matemáticos también se toparon de forma ocasional con la

cuestión de la imposibilidad dentro del contexto de varios problemas fundamentales de aritmética, geometría y álgebra. Supuestamente, alrededor del 550 a. C. los pitagóricos se toparon por primera vez con la «irracionalidad» (que en su origen significaba tan solo «no fracción» o «no razón», en el sentido de no ratio, pero no implicaba «fuera de la razón», como cabría pensar hoy en día) de números como $\sqrt{2}$, que no podía expresarse como una fracción de dos enteros[5]. Cuenta la leyenda que los pitagóricos consideraron tan escandalosa esta revelación que ahogaron a su descubridor, Hípaso de Metaponto, por los problemas que les causó. Esto da una ligera idea sobre el tipo de operaciones y cuestiones que no tienen respuesta a partir de una serie particular de reglas.

Durante el primer cuarto del siglo XIX, el joven matemático noruego Henrik Abel demostró que no hay manera de resolver de forma explícita una ecuación algebraica general de quinto grado ($Ax^5 + Bx^4 + Cx^3 + Dx^2 + Ex + F = 0$, con $A, B..., F$ constantes arbitrarias) aplicando a sus coeficientes operaciones aritméticas ordinarias, como la suma, la resta, la multiplicación, la división y la extracción de raíces del grado adecuado (hasta la quinta en este caso)[6]. El premio equivalente al Nobel de matemáticas lleva hoy el nombre de este físico. A diferencia de las ecuaciones de segundo, tercer o cuarto grado, las de quinto grado no se pueden resolver con una fórmula exacta[7]. Tan solo unos años después, en 1837, se demostró con rigor de varias maneras que un ángulo de 60º no se puede dividir en tres partes iguales trazando tan solo rectas o arcos de circunferencia con escuadra y compás. Estos casos revelaron por primera vez a quienes los estudiaron del modo correcto algunos indicios de las limitaciones de determinados sistemas axiomáticos.

En 1899 David Hilbert, el matemático más brillante del momento, emprendió un plan sistemático para apoyar las matemáticas sobre un plano axiomático formal[8], un programa que expuso en su totalidad en el ensayo titulado «Sobre el infinito» de 1925[9]. Hilbert creía que sería posible determinar los axiomas en los que se fundamenta cada parte de las matemáticas (y, por tanto, todas ellas en su conjunto), demostrar la coherencia interna

de esos axiomas y después evidenciar que el sistema resultante de enunciados y deducciones surgido de esos axiomas es completo y, al mismo tiempo, decidible.

Más en concreto, un sistema es «coherente» cuando no se puede demostrar que un enunciado *E* y su negación, *–E*, son teoremas verdaderos (por ejemplo, demostrando que un número es al mismo tiempo par e impar).

Y es «completo» si para cada enunciado *E* que se pueda construir en el marco de su lenguaje se puede demostrar que es cierto, o bien el propio enunciado *E*, o bien su negación.

Por último, el sistema es «decidible» si se puede demostrar que es verdadero o falso cualquier enunciado E construible con ese lenguaje. Por tanto, si un sistema es decidible también tiene que ser completo.

La visión formalista que tenía Hilbert de las matemáticas consistía en una densa red de deducciones que se desplegaba con impecables conexiones lógicas a partir de los axiomas fundamentales. De hecho, las matemáticas se definieron como el conjunto de todas esas deducciones. Hilbert se propuso completar esa formalización de las matemáticas con ayuda de otros colegas, y creyó que en última instancia sería posible ampliar su alcance para incluir ciencias que se fundamentan en las matemáticas aplicadas como, por ejemplo, la física [10]. Él partió de la geometría euclídea y logró apoyarla sobre una base axiomática rigurosa. Después, su plan imaginó el reforzamiento del sistema con la incorporación de axiomas adicionales que evidenciaran a cada paso el mantenimiento de la coherencia y la decidibilidad, hasta que el sistema acabara siendo lo bastante extenso como para abarcar toda la aritmética. El criterio esencial para acreditar estos sistemas era siempre el mismo: demostrar la coherencia.

Hilbert puso en marcha su plan con gran confianza y convencido de que solo era cuestión de tiempo que todas las matemáticas quedaran atrapadas dentro de su red formalista. Esta confianza se refleja en el epitafio de su tumba, que es una cita del discurso que pronunció ante la Sociedad Alemana de Naturalistas y Físicos el 8 de septiembre de 1930 con motivo de una conferencia sobre epistemología de las ciencias exactas:

Wir müssen wissen
Wir werden wissen,

es decir, «Debemos saber. Sabremos». Este convencimiento incondicional se manifestó en otras declaraciones no matemáticas de Hilbert. Hablando sobre el proceso judicial de Galileo y su incapacidad para defender sus convicciones científicas, Hilbert subrayó que el científico pisano «no era estúpido […]. Solo un estúpido podría creer que la verdad científica necesita recurrir al martirio; este tal vez sea necesario para la religión, pero los resultados científicos se acaban demostrando a su debido tiempo[11]». Pero, mira por dónde, el joven Kurt Gödel puso el mundo patas arriba algo más tarde, apenas un día después del convencido discurso de Hilbert, durante una mesa redonda al margen de la conferencia de Königsberg.

Gödel había completado uno de los primeros pasos del plan de Hilbert como parte de su tesis doctoral, en la que demostró la coherencia y completitud del primer orden lógico (aunque más tarde Alonzo Church y Alan Turing revelarían que no es decidible). En 1929 el matemático polaco Mojżesz Presburger demostró que la aritmética de Peano es coherente, decidible y completa si no se incluye la multiplicación. En su caso se habla de la aritmética de Presburger. En 1930 el noruego Thoralf Skolem demostró lo mismo sobre la aritmética de Peano, aunque dejando fuera la adición. Pero los siguientes pasos que acometió Gödel fueron los que le valieron la fama de ser el lógico más grande que ha existido desde Aristóteles. Lejos de expandir el plan de Hilbert para alcanzar su objetivo principal (una demostración de la completitud de la aritmética), Gödel demostró lo contrario: que cualquier sistema coherente lo bastante rico como para contener toda la aritmética tiene que ser incompleto e indecidible. Esto pilló a casi todo el mundo por sorpresa y, de hecho, acabó de un solo golpe con el plan de Hilbert (y también con el de Whitehead y Russell).

Gödel provocó reacciones optimistas y pesimistas. Las optimistas, como la de Freeman Dyson, consideran sus resultados como la garantía de la naturaleza interminable de la investigación humana. Contemplan la indagación científica como una parte tan esencial del espíritu humano que, si se completara, produciría un

efecto desmotivador desastroso en nosotros. En cambio, las interpretaciones pesimistas, como la de John Lucas [12], consideran el resultado de Gödel como la confirmación de la imposibilidad de la mente humana para conocer todos los secretos de la naturaleza (tal vez incluso la mayoría de ellos).

Sin embargo, otros, como Stanley Jaki [13], contemplan esta barrera de incompletitud con optimismo, ya que demostraría que la mente humana no puede ser superada por una máquina calculadora. Otros usos, más bien informales, de los resultados de Gödel se revelan divertidos. Por ejemplo, si una religión es un sistema de pensamiento que contiene afirmaciones no demostrables que deben aceptarse por la fe, entonces las matemáticas no solo son una religión, sino que también es la única capaz de demostrar que es una religión.

Es interesante observar que ninguna de las demostraciones de «imposibilidad» previas, como la de Abel, dio lugar a estas consideraciones generales sobre la capacidad cognitiva última y definitiva de la especie humana. En el caso de Abel, por ejemplo, se descarta una fórmula general para resolver ecuaciones de grado fijo superior a 4, pero se consigue a partir de los coeficientes de la ecuación mediante operaciones elementales (es decir, suma, etc.) y por extracción de raíces. Pero si se permiten herramientas más poderosas, se puede encontrar una fórmula. El discurso de Gödel, sin embargo, es más radical.

Gödel fue un hombre, cuando menos, singular. Tal como escribió el poeta John Dryden en 1861, «los grandes genios son sin duda aliados íntimos de la locura. Y finas son las particiones que separan sus fronteras [14]».

Cuando he preguntado a algunos de los contemporáneos de Gödel del Instituto de Estudios Avanzados de Princeton si habían llegado a conocerlo, siempre me han respondido: «nadie conocía a Gödel». Un físico famoso me dijo que, cuando llegó a ese instituto siendo muy joven, estaba, como es natural, ansioso por hablar con Gödel sobre sus teoremas de incompletitud y otras cuestiones relacionadas con la mecánica cuántica. Lo llamó a través de la línea telefónica interna y lo sorprendió gratamente que lo pusieran

directamente con él, sin la mediación de ninguna secretaria o asistente que lo parapetara. Gödel le ofreció un encuentro en su oficina. El joven físico se quedó entusiasmado con la idea, pero al llegar la hora convenida no encontró a nadie. Pensó que Gödel se habría retrasado por algún compromiso importante. Al día siguiente asistió al té de bienvenida que ofrecía el instituto a los recién llegados y vio a Gödel sentado solo en un rincón. Se acercó a él, se presentó y le explicó que había ido a su oficina a la hora acordada, pero que se figuraba que habría tenido que irse por algún asunto más importante. «Al contrario», respondió Gödel, «he aprendido que solo concertando una cita puedo estar seguro de que no me encontraré con nadie».

Es útil concretar las afirmaciones precisas que subyacen a la deducción de incompletitud de Gödel. El teorema de Gödel establece que si un sistema formal es

1. *especificado de manera recursiva*,
2. *lo bastante grande para incluir la aritmética*,
3. *coherente*,

entonces se deduce, precisamente, que es incompleto.

La condición 1 significa que debe haber una relación clara y accesible de axiomas como, por ejemplo, en la aritmética de primer orden de Peano.

La condición 2 significa que el sistema formal consigue definir los números naturales junto con sus operaciones de suma y multiplicación (¡ambas!) y demostrar ciertos hechos elementales sobre ellos. Esto se da, de nuevo, con la aritmética de Peano de primer orden.

La estructura de la aritmética tiene un papel central en la demostración del teorema de Gödel. Gödel utilizó ciertas propiedades especiales de los números, como sus factorizaciones y, tal como reveló Gauss, el hecho de que cualquier número se pueda factorizar de una manera única como el producto de divisores primos (por ejemplo, $130 = 2 \times 5 \times 13$), para establecer una correspondencia crucial entre enunciados matemáticos y enunciados sobre matemáticas. Estos últimos se denominan en ocasiones enunciados metamatemáticos.

Gödel llegó a su resultado asignando a cada afirmación lógica un

número de código, es decir, un número natural que la representa. Para ello se basa en el teorema de la descomposición en factores primos. Primero, asigna un número de código a los distintos símbolos implicados: por ejemplo, 3 al símbolo cero (0), o 19 al símbolo de la relación de igualdad (=). En este punto, una fórmula viene codificada por el número que se obtiene tomando los diversos factores primos 2, 3, 5, 7… y elevándolos a los números que se han asignado a los símbolos de la fórmula, en el orden en que aparecen. Así, «0 = 0» está numerado como $2^3 \times 3^{19} \times 5^3$. De esta manera, las afirmaciones sobre los números se convierten en números de por sí, y permiten expresar como enunciados sobre los números naturales las afirmaciones que se refieren, sin embargo, a razonamientos sobre los números naturales. En un sistema dado *E* se puede formular, por tanto, como una proposición sobre los números, una afirmación que habla de sí misma y dice: «No soy demostrable mediante *E*». Lo que conlleva la consecuencia de que esa afirmación es demostrable si y solo si no lo es (otro recordatorio de la paradoja del mentiroso). ¡Un movimiento de jiujitsu matemático!

De modo que la aritmética de primer orden de Peano no es completa ni decidible.

Una vez más resulta instructivo ver que es posible no reunir esas condiciones. Si se elige una teoría coherente en cuanto a referencias a (y relaciones entre) tan solo los 10 primeros números naturales, es decir, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, y se suman y multiplican dentro de un sistema de módulo 10 (es decir, cada vez que se pasa de 9 se vuelve a empezar desde el 0, por ejemplo, 5 + 6, que es 11, se convierte de nuevo en 1), entonces esta «miniaritmética» es completa. Su universo es, de hecho, finito y escapa a la condición 2 del teorema de Gödel.

Un resultado mucho más sorprendente lo ofrece la teoría de primer orden de los números reales con sus respectivas operaciones de adición y multiplicación, que es un mundo aparentemente mucho más complicado que el de los números naturales. Sin embargo, esta teoría es completa y decidible, tal como constata un teorema de Alfred Tarski, demostrado en medio de la última guerra mundial. Dado que la geometría euclídea se expresa mediante los números reales y sus operaciones, se puede deducir que también es

completa. Cabría objetar que no hay nada mágico y, por lo tanto, impredecible en la naturaleza euclídea, plana, de la geometría clásica. Pues bien, hasta las geometrías no euclídeas de las superficies curvas resultan ser completas.

Pero el asunto no se zanja con un teorema abstracto de completitud y decidibilidad, como el que acabamos de enunciar para los números y la geometría reales. Cualquier procedimiento para diferenciar lo verdadero de lo falso, es decir, un algoritmo de decisión para la teoría de números reales de primer orden, puede requerir en la práctica un largo y tortuoso camino de verificación. En efecto, un teorema de 1974 de Fischer y Rabin sostiene que cualquier algoritmo para reconocer enunciados verdaderos para los números reales con adición y multiplicación (e incluso para números reales solo con adición) tarda en responder, en el peor de los casos, tiempos al menos exponenciales respecto de la longitud del enunciado que se ha de analizar y, por lo tanto y en definitiva, amenaza con esperas prohibitivas.

Peor aún es la aritmética aparentemente inofensiva de Presburger, la única que es completa y decidible incluyendo solo la adición. Sin embargo, cualquier algoritmo que decida sus enunciados requiere, en el peor de los casos, un tiempo doblemente exponencial respecto de la longitud del enunciado en cuestión. Por lo tanto, anuncia esperas muchísimo más prolongadas incluso que en el caso de los números reales. Así lo demostraron Fischer y Rabin.

Como ya se ha dicho, muchas teorías físicas se basan en números reales o en números complejos, y existe una riqueza sorprendente de fenómenos que se pueden formalizar en el ámbito abstracto de estos conjuntos numéricos con sus operaciones. Todo esto infunde la esperanza de que estas teorías físicas puedan evitarse la vergüenza de la indecidibilidad. Tarski, con la ayuda de sus colaboradores, entre los que se encontraba su alumna Wanda Szmielew, señaló que muchos sistemas matemáticos utilizados en física, como la teoría de retículos, la geometría proyectiva y la teoría de los grupos abelianos (así llamada en honor a Abel), son decidibles. Otros, sin embargo, como la teoría de grupos no abelianos, no lo son [15].

Así termina un capítulo extraordinario de la historia del

desarrollo de las matemáticas, que tiene ramificaciones en física, ciencias de la computación y el poder de la mente humana en relación con sistemas axiomáticos como los ordenadores.

9. Por qué son tan comunes los unos y los doses

El presidente se bebía una botella de champán cuando ganábamos. Cuando perdíamos se bebía dos y pensaba que habíamos ganado.

Bobby Robson [1].

Cuando usamos los números con cualquier finalidad (ya sea para contar, enseñar, en el banco, para realizar encuestas, en ingeniería, para calcular resultados de equipos deportivos o para viajar) ocurre algo curioso con ellos, sobre todo con los más pequeños, el 1 y el 2, en lo que tal vez no haya reparado usted. Este detalle fascinante de las matemáticas simples relacionado con los números pequeños fue observado por primera vez por el astrónomo estadounidense Simon Newcomb [2] en 1881, y redescubierto después en 1938 por el ingeniero Frank Benford, quien escribió sobre él y acabó dando su nombre a lo que ahora se conoce como ley de Benford.

Lo que notaron Newcomb y Benford fue que en muchos conjuntos de números que parecen confluir de forma aleatoria, como áreas de lagos, puntos de béisbol, potencias de 2, números en el interior de revistas, posiciones de estrellas, listas de precios, constantes físicas o registros de contabilidad, el primer dígito sigue una distribución de probabilidad muy particular y bastante precisa con independencia de la unidad de medida utilizada [3]. Con «primer dígito» de un número real positivo nos referimos a la primera cifra significativa que aparece en la representación decimal de ese número, por ejemplo, el 1 en 0.013 82. Véase la tabla 9.1.

Cabría esperar que los nueve dígitos que van del uno al nueve

aparecieran con la misma probabilidad en la primera posición, es decir, con una probabilidad aproximada de $1/9 = 0.111$ (de tal manera que la suma de las nueve probabilidades se acercara a 1 como buena aproximación). Sin embargo, Newcomb y Benford descubrieron que los primeros dígitos d, sin contar el cero, tienden a seguir otra ley simple de frecuencia en el caso de listas de un tamaño considerable que arroja la probabilidad $P(d)$ para el dígito d que ocupa la primera posición después de la marca decimal:

$$P(d) = \log_{10}(d + 1) - \log_{10}(d) = \log_{10}[1 + 1/d],$$
$$d = 1, 2, 3\ldots, 9.$$

Esta regla predice lo siguiente (redondeando al segundo decimal) para la probabilidad de que los diferentes dígitos aparezcan en primer lugar (tabla 9.1):

$$P(1) = 0.30,\ P(2) = 0.18,\ P(3) = 0.12,$$
$$P(4) = 0.10,\ P(5) = 0.08,\ P(6) = 0.07,$$
$$P(7) = 0.06,\ P(8) = 0.05,\ P(9) = 0.05.$$

El número 1 es el que presenta más probabilidad de aparición en primer lugar, con una frecuencia de 0.30, mucho mayor que la ocurrencia igual de probable de $1/9 = 0.11$. Hay varias maneras bastante sofisticadas de llegar a la fórmula para $P(d)$, la cual nos dice que las probabilidades de aparición de los diferentes dígitos están uniformemente distribuidas sobre una escala logarítmica (figura 9.1). Además, si cambiamos la base aritmética de 10 a b, por decir algo, entonces lo único que cambia en la fórmula Newcomb-Benford es que la base del logaritmo en la fórmula para $P(d)$ pasa de 10 a b. De modo que en el caso de la aritmética binaria ($b = 2$), las cosas son muy simples. Los únicos números del sistema binario que no empiezan por 1 son el cero y el conjunto vacío, por tanto, $P(1) = \log_2(2) = 1$, y en términos estadísticos, en el límite, la probabilidad de que el primer dígito binario valga 1 es del 100%.

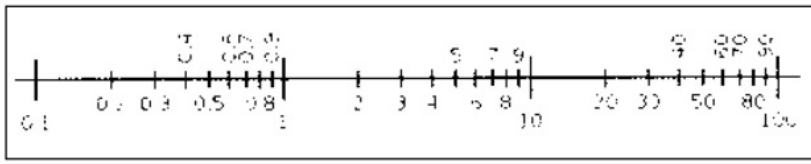


Figura 9.1. Línea logarítmica de números. Al elegir al azar un punto x de manera uniforme sobre esta línea, alrededor del 30% de las veces el primer dígito del número elegido será 1.

Conviene comprender por qué hay una tendencia mayor hacia los dígitos más bajos, sobre todo el 1 y el 2. Pensemos en la posibilidad de que el primer dígito sea 1 a medida que crece la lista de todos los números, como 1, 2, 3, 4…, 9, 10, 11…, 19, 20…, 99, 100, 101, etc. Consideremos tan solo los dos primeros números, 1 y 2; en este caso, la probabilidad de que el 1 sea el primer dígito en aparecer es claramente de tan solo 1/2. Si incluimos todos los números hasta el 9, entonces la probabilidad $P(1)$ cae hasta 1/9. Cuando incorporamos el siguiente número, el 10, aumenta hasta 1/5 porque ahora son dos los números de todo el conjunto que empiezan por 1 (1 y 10). Si ahora incluimos 11, 12, 13, 14, 15, 16, 17, 18, 19, de repente $P(1)$ se dispara a 11/19. Continuemos hasta el número 99 y ya no habrá más números que comiencen por 1, así que $P(1)$ vuelve a reducirse: después de 99 números cae hasta $P(1) = 11/99$. Al llegar al 100, la probabilidad vuelve a aumentar hasta el 199, porque todos los números que van del 100 al 199 empiezan por 1.

Vemos, pues, que la probabilidad $P(1)$ de que el primer dígito sea 1 aumenta y desciende a partir de los números 10, 100, 1000…, y continúa con un patrón de dientes de sierra. La ley de Newcomb-Benford no es más que el promedio de esos aumentos y descensos en la gráfica de dientes de sierra de $P(1)$, que oscila en torno al valor promedio. La media arroja un valor aproximado de 0.30, tal como predice la fórmula para $P(d)$.

La ubicuidad de la ley de Newcomb-Benford es muy llamativa y se ha usado como herramienta para detectar declaraciones fiscales sospechosas (los números tienden a no seguir esta ley cuando se

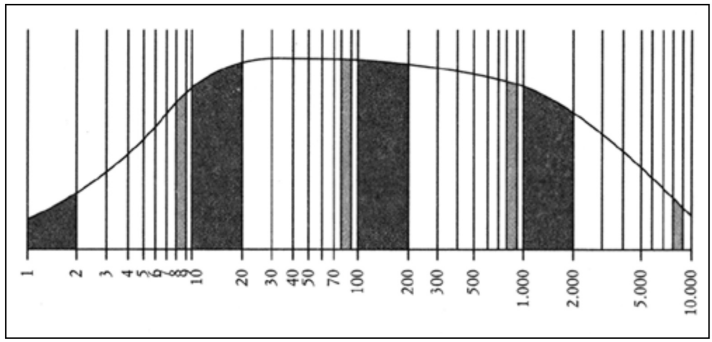
inventan o se crean de manera artificial mediante un generador aleatorio de números, en lugar de hallarse de forma natural mediante la suma de muchas entradas de contabilidad). El introductor de esta idea en el ámbito de la contabilidad fue Mark Nigrini, un doctorando de la Universidad de Cincinnati que la usó en 1992 [4], y desde entonces se ha empleado con bastante eficacia para detectar datos fraudulentos. El inspector jefe de la oficina fiscal del distrito de Brooklyn probó el método de Nigrini en retrospectiva con siete casos de fraude fiscal ya detectados con anterioridad y consiguió localizarlos todos de manera independiente. Hasta se empleó con la declaración de la renta de Bill Clinton, pero no hubo nada sospechoso. El punto débil de estos análisis es que si los números se redondean de manera sistemática, los datos brutos se sesgan y la prueba deja de funcionar.

Aunque la ley de Newcomb-Benford es ubicua, no es totalmente universal: no se trata de una ley de la naturaleza en un sentido físico [5]. La distribución de la estatura o el peso de una persona, de datos de coeficiente intelectual, números de teléfono, números de direcciones o las combinaciones ganadoras de la lotería no parecen seguir la ley de Newcomb-Benford. ¿Cuáles son las condiciones necesarias para que esta ley describa la distribución de los primeros dígitos dentro de un conjunto de números?

Los datos utilizados deben proceder únicamente de cantidades del mismo tipo: no se pueden sumar áreas de lagos con números de datos fiscales. No debe haber ningún límite prefijado para los números más grandes o más pequeños de la colección, como suele ocurrir con la numeración de viviendas, y los números no deben asignarse mediante ningún sistema especial, como sucede, por ejemplo, con los códigos postales, números de viviendas, números de pasaporte o números de teléfono.

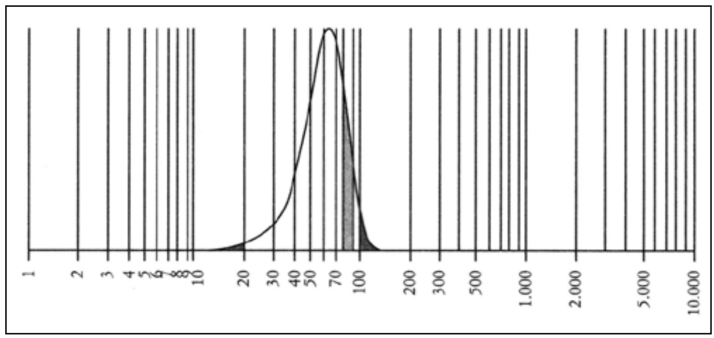
La distribución de las frecuencias debe ser bastante regular, sin grandes picos alrededor de números particulares. Sin embargo, lo más importante es que debe abarcar una gama amplia de cantidades (decenas, centenas y millares) y que la distribución de frecuencias sea amplia y plana en lugar de tener picos apretados en torno a un valor medio determinado; véanse las figuras 9.2a y 9.2b, donde se ilustran de forma gráfica ambas condiciones. Cuando se traza la distribución de probabilidad es porque se quiere que el área bajo la

curva para algún intervalo de resultados esté determinada en gran parte por la anchura de la distribución, y no por su altura (como en la figura 9.2a). Si la distribución es relativamente estrecha y picuda, como en la frecuencia de la altura de los adultos de la figura 9.2b, entonces el primer dígito de esas alturas no seguirá la ley de Newcomb-Benford. Aun así, esto no explica la ubicuidad de la ley de Newcomb-Benford porque no aclara la ubicuidad de los conjuntos de datos con estas características requeridas en los mundos natural y humano.



.

Figura 9.2a. Distribución de probabilidad amplia del logaritmo de una variable, plasmada sobre una escala logarítmica. Esta arrojará como resultado la distribución de Newcomb-Benford de los primeros números.



.

Figura 9.2b. Distribución de probabilidad estrecha del logaritmo de una variable, plasmada sobre una escala logarítmica. Esto no arrojará como resultado la distribución de Newcomb-Benford de los primeros números.

Newcomb reparó por primera vez en este fenómeno de los números más bajos cuando se dio cuenta de lo gastadas que estaban las páginas de su libro de logaritmos que contienen los dos primeros números. Es una particularidad sorprendente y revela que incluso en la actualidad sigue habiendo algo especial en los números más pequeños. Se usan mucho. Recordemos que, como hemos visto, los dígitos elegidos al azar deben tener el 1 y el 2 como primer dígito después de la marca decimal en el 11% de los casos. Newcomb y Benford predicen que el 1 aparecerá el 30% de las veces, mientras que el 2 lo hará el 18%. En cambio, los números mayores que 4 aparecen menos que la expectativa aleatoria del 11% de los primeros dígitos. Haga la prueba con algunos números.

10. ¿Qué son las matemáticas?

Los biólogos creen ser bioquímicos,
los bioquímicos creen ser fisicoquímicos,
los fisicoquímicos creen ser físicos,
los físicos creen ser Dioses,
y Dios cree ser matemático.

Anónimo [1].

Hemos visto muchas matemáticas en este libro, pero ¿qué son las matemáticas? Si preguntamos a un historiador o químico qué es la historia o la química, serán capaces de responder con toda claridad. Los matemáticos siempre han tenido problemas con esta pregunta, y a la mayoría de ellos no le preocupa lo más mínimo mientras desempeña su labor matemática, aunque pueda conversar sobre el asunto los fines de semana o con quienes no se dedican a ellas. Salvo en el caso de unas pocas excepciones, se ha convertido en el dominio de un pequeño grupo de filósofos de las matemáticas, y hay cuatro opciones para elegir.

La primera respuesta se denomina platonismo matemático y forma parte de la idea platónica de que todo lo que vemos es una manifestación parcial o imperfecta de arquetipos ideales o perfectos que existen en algún otro mundo. De acuerdo con esta concepción, los objetos matemáticos —como conjuntos, curvas, exponentes y números— tienen una existencia real, por eso el ser humano los descubre en lugar de inventarlos. A los matemáticos les gusta sentirse así cuando hacen matemáticas. Si encuentran un método nuevo para resolver un viejo problema les gusta pensar que lo han descubierto gracias a su ingenio y esfuerzo. Cuando se profundiza en esta filosofía afloran algunos aspectos extraños. Admite un

mundo de ejemplos ideales, de construcciones artificiales que hacen referencia a las cosas cotidianas, pero con el que nosotros, pobres mortales confinados en el fondo de la caverna platónica, no podemos tener ninguna interacción en absoluto. Aristóteles no aceptaba esta cosmovisión porque pensaba que la naturaleza de las cosas deriva de las formas que hay dentro de ellas (*informar* es «dar forma y sustancia a una cosa», de ahí el sentido actual de la expresión «informar a alguien»). Y él argumentó que esta visión platónica conduce a una regresión al infinito:

> Si todos los hombres son iguales porque tienen algo en común con el Hombre, que es el arquetipo ideal y eterno, ¿cómo se explica el hecho de que el hombre y el Hombre se parezcan entre sí sin asumir la existencia de otro arquetipo? ¿Y no nos llevaría este mismo razonamiento a considerar necesarios un tercer, un cuarto y un quinto arquetipo y así sucesivamente en una regresión hacia más y más mundos ideales?

Otro problema surge al asumir que nuestros triángulos y líneas paralelas imperfectos son reflejos parciales de su arquetipo real y perfecto. ¿Son ejemplos particulares imperfectos de un arquetipo perfecto o son ejemplos perfectos de un arquetipo imperfecto? Cualquier objeto matemático que pueda existir existe en el «paraíso» platónico de los arquetipos ideales, a la espera de que lo descubramos y lo adoptemos.

La filosofía platónica floreció en la tradición cristiana, donde Dios se consideraba el creador todopoderoso de las leyes del universo. Las matemáticas formaban parte de la mente de Dios, parte de la verdad última sobre el universo: la senda que más nos permite acercarnos a esos cimientos de la verdad. Antes de que los matemáticos empezaran a encontrar geometrías diferentes a la de Euclides y lógicas distintas de la de Aristóteles, las matemáticas se correspondían con la forma del universo. Cualquier crítica que cuestionara la esperanza de que los teólogos humanos accedieran a las verdades últimas del universo se desmontaba apelando a la geometría euclídea como vía de acceso a una parte de la realidad absoluta. En la práctica, los matemáticos tienden a trabajar como si esto fuera la verdad, aunque lo más probable es que no lo defiendan como filosofía de trabajo. El platonismo es una forma de realismo

sobre las matemáticas que de entrada parece simple y llana, pero que resulta ser muy compleja cuando se le intenta dar sentido [2]. Entre los matemáticos platónicos más célebres se cuentan: Frege, Gödel, Cantor, Willard van Quine y Roger Penrose.

En resumen: el platonismo matemático sostiene que los objetos matemáticos (como los números 1 y 2 y la fórmula 1 + 1 = 2) existen y que son objetos abstractos (es decir, no pueden tener efectos tangibles). Además, existen con independencia de los pensamientos y del discurso que pueda mantener sobre ellos cualquier ser inteligente. Forman parte de la realidad al margen de lo que pensemos sobre ella. Desde este punto de vista, es de esperar que cualquier inteligencia extraterrestre descubra las mismas matemáticas que tenemos nosotros, aunque las exprese en un lenguaje propio, diferente del nuestro, y utilice, tal vez, una aritmética en base 8, si es que tienen 8 «dedos». De hecho, todo el programa SETI (del inglés *Search for Extraterrestrial Intelligence*, «búsqueda de inteligencia extraterrestre») que se ha desarrollado dentro de la astronomía durante décadas con radiotelescopios se ha basado en el convencimiento de que existen verdades físicas y matemáticas universales que las personas hemos descubierto y que cualquier inteligencia extraterrestre ha tenido que desentrañar también si está capacitada para recibir y enviar señales.

La primera alternativa al platonismo matemático medieval europeo fue el desarrollo de una filosofía llamada formalismo matemático y con la que nos hemos topado en algunos capítulos de este libro. Emergió como respuesta a la proliferación de diferentes geometrías y lógicas que evidenciaban que las matemáticas son una verdad única y singular. Matemáticos como Frege, Peano, Russell, Whitehead y Hilbert aspiraron a inferir todas las matemáticas a partir de un conjunto único de axiomas evidentes y coherentes entre sí. Para ellos, la verdad matemática, o su «existencia», significaba simplemente coherencia interna. No les interesaba en absoluto establecer una correspondencia entre las matemáticas y el mundo físico real. Las matemáticas eran un «juego», como lo son el ajedrez, el go o el GoL (*Game of Life*, «Juego de la Vida», de John Conway, un autómata celular basado en los fundamentos del crecimiento celular en cultivos), dotado de reglas y unas posiciones específicas de partida. Ya hemos visto cómo se aplicó este procedimiento en la

aritmética. Las matemáticas se reducían al conjunto de posiciones de partida sobre el tablero de juego desde la casilla inicial, y no significaban nada fuera de ese tablero de juego.

Hilbert consideraba posible mostrar que todas posiciones permitidas sobre el tablero (teoremas o fórmulas) se pueden alcanzar a partir de la posición inicial mediante una cantidad finita de «movimientos», lo que permitiría determinar si cualquier posición dada sobre el tablero se puede alcanzar o no aplicando las reglas a partir de la posición de partida. Asimismo, había que comprobar si no surgía ninguna incoherencia, de tal modo que las reglas no permitieran deducir que tanto A como su opuesto son verdad, ya que con incoherencias así siempre se podría demostrar que 1 = 2. Hilbert estaba convencido de la viabilidad de su plan, pero estaba muy equivocado. Como hemos visto, Gödel demostró la imposibilidad de llevarlo a cabo. Cuando un sistema lógico alcanza la complejidad de la aritmética, ya no se pueden demostrar ni refutar las aspiraciones de Hilbert. Es como si nos muestran una configuración válida de las piezas de ajedrez sobre el tablero, pero ninguna configuración previa conduce hasta ella. El ajedrez, como la verdad matemática, está más allá de cualquier demostración.

Muchos especialistas en matemáticas, sobre todo en matemáticas puras, son formalistas. Si no tienen ningún interés en aplicar sus estructuras matemáticas a la ciencia o a otros aspectos del mundo real, entonces son felices explorando la vasta complejidad de las estructuras matemáticas, como los participantes en el juego de los abalorios en la novela homónima de Hermann Hesse. Para los formalistas, las matemáticas son una invención, no un descubrimiento. Establecemos reglas y luego indagamos en las consecuencias de esos conjuntos de axiomas. Lo que se infiere a partir de ellos no existe más que en nuestra mente y en nuestros cuadernos de notas y pizarras.

La tercera variante de la filosofía matemática es el constructivismo, dentro del cual también puede incluirse el intuicionismo. Los grandes defensores decimonónicos de esta corriente fueron el alemán Leopold Kronecker y el matemático catatónico holandés Luitzen Brouwer. Como hemos visto, la concepción de ambos se oponía a la de Cantor, cuyos infinitos necesitaban residir en un paraíso platónico. El constructivismo

reducía el alcance de la verdad matemática. Consideraba «verdaderos» únicamente aquellos enunciados matemáticos que pudieran deducirse progresivamente a partir de los axiomas y usando las reglas mediante una cantidad finita de pasos. Era una concepción más reducida que las matemáticas convencionales de los platónicos porque prohibía los métodos tradicionales de demostración no constructivistas, es decir, las demostraciones de la existencia de un objeto matemático al margen de su construcción y representación explícitas paso a paso. Los constructivistas querían proscribir estas «demostraciones» porque temían que permitieran que alguna contradicción lógica se infiltrara en las matemáticas y las derrumbara, del mismo modo que los griegos temieron admitir el 0 en sus matemáticas (¿cómo va a ser algo la nada?). Los constructivistas no aceptarían demostraciones que supusieran algo como verdadero y después permitieran inferir una contradicción a partir de esa suposición, lo que evidenciaba que no podía ser verdad. Veamos un ejemplo precioso de Euclides, del siglo IV a. C., expresado en un lenguaje matemático moderno.

Para demostrar que hay una infinidad de números primos, Euclides parte del supuesto de que esa afirmación es falsa, de modo que los números primos quedarían reducidos a una cantidad finita $p_1 \ldots, p_n$. Si se multiplican todos ellos entre sí y se les suma 1 obtenemos $Q = p_1 \times \ldots \times p_n + 1$. Con respecto a Q caben dos posibilidades:

- que Q sea un número primo, aunque claramente distinto de $p_1 \ldots, p_n$, por lo que sería un nuevo número primo;
- que Q sea un número compuesto, y en ese caso tiene un factor primo q, que también sería distinto de los enumerados puesto que no puede figurar entre $p_1 \ldots, p_n$, porque, de lo contrario, sería un divisor tanto del producto $p_1 \times \ldots \times p_n$ como de la diferencia $Q - p_1 \times \ldots \times p_n$, que es 1.

Se demuestra así que, si los números primos son finitos (y, por tanto, se da por supuesta la existencia de un número primo mayor que todos los demás), entonces se llega a una contradicción lógica.

Veamos ahora ejemplos de las dos posibilidades a las que llega la demostración de Euclides.

- Supongamos que creemos que 2, 3, 5 son todos los números primos que hay y, por tanto, que 5 es el número primo más elevado. Repitamos entonces la operación $2 \times 3 \times 5 + 1 = 31$, lo que da como resultado un nuevo número primo mayor que 5.
- Supongamos, en cambio, que creemos que 2, 3, 5, 7, 11 y 13 son todos los números primos que existen y, por tanto, que 13 es el primo más alto. Entonces $2 \times 3 \times 5 \times 7 \times 11 \times 13 + 1 = 30\,031$ que, aunque no es primo, tiene dos factores primos, 59 y 509, que son dos números primos nuevos, es decir, distintos de los iniciales.

Nótese que el razonamiento de Euclides se podría convertir fácilmente en constructivista: a partir de cualquier conjunto finito de números primos, se crea uno nuevo siguiendo sin más la estrategia propuesta.

El constructivismo nunca gozó de popularidad entre los matemáticos, a pesar de ser un ejercicio perfectamente válido desde el punto de vista formalista, porque su sistema de reglas admite menos cosas como verdaderas, y desarrollos matemáticos apasionantes de muchos tipos, como los infinitos de Cantor, quedan fuera de sus competencias. Por decirlo de alguna manera, es como hacer matemáticas con un brazo atado a la espalda. Sin embargo, a mediados de la década de 1950, cuando la programación informática se convirtió en una aplicación importante del enfoque constructivista de las matemáticas, esta filosofía encontró un nicho más natural. Aun así, al igual que el formalismo, no tiene ningún interés por relacionar las matemáticas con el mundo físico, y las contempla en su totalidad como una construcción del intelecto humano.

Una rama moderna de este planteamiento recibe el nombre de constructivismo social y considera las matemáticas como una construcción de la sociedad. El conocimiento objetivo que posee en la actualidad se ha derivado y desarrollado a partir del conocimiento colectivo y las interacciones sociales. De ahí que las matemáticas no sean un conjunto de teoremas o un texto incuestionable, sino algo parecido al conjunto de conocimientos legales de un país o la interpretación de su constitución, o al estudio de la cosmología. Es algo que existe en diferentes partes y en

muchas mentes [3], y que solo pasa a formar parte del corpus colectivo de las matemáticas cuando es aceptado por muchos matemáticos.

La última filosofía de las matemáticas que consideraremos aquí es el estructuralismo, y es la que me parece más útil. Pone el énfasis en que las matemáticas son un conjunto de patrones o relaciones entre cosas, más que sobre las cosas de por sí, como los números. Imagine que alguien le dice que la ubicuidad de las matemáticas le resulta desconcertante: describe casi todo lo que hay en el mundo. Funciona especialmente bien dentro de la física, la astronomía y la ingeniería, aunque algo peor en las ciencias sociales, la psicología y las humanidades. Creo que la respuesta correcta es que las matemáticas son el conjunto de todos los patrones posibles: algunos muy evidentes, como las órbitas de los planetas, las simetrías o las pautas de interacción entre partículas elementales de la materia; otros menos inmediatos, como los números complejos; y otros, como los infinitos de orden superior de Cantor, que aún no han encontrado un uso dentro de la naturaleza, al menos hasta ahora.

Por tanto, es inevitable que las matemáticas «funcionen» y describan las leyes —y sus resultados— de las ciencias físicas. No podríamos existir si no hubiera patrones en el universo. De hecho, el universo sería un vacío absoluto sin ninguna estructura. Si hay patrones tiene que haber unas matemáticas que los describan [4]. Sin embargo, esto no despeja el misterio de lo que el físico Eugene Wigner denominó «la irrazonable eficacia de las matemáticas en las ciencias naturales [5]». Porque, aunque es obvio que las matemáticas tienen que describir los patrones del mundo natural, puesto que describen todos los patrones posibles, el misterio es por qué un número tan reducido de patrones relativamente simples es una herramienta tan versátil y poderosa para explicar y comprender el universo. Sabemos que el universo podría haber sido mucho más complejo (en realidad mucho más de lo que nuestras mentes alcanzarían a comprender incluso hoy). Son interesantes las palabras que dedicó Paul Dirac a esta cuestión:

> El matemático juega un juego cuyas reglas ha inventado él mismo, mientras que el físico practica un juego cuyas reglas vienen dadas por la naturaleza; sin embargo, a medida que pasa el tiempo se vuelve cada vez más evidente que las reglas que el matemático encuentra

interesantes son las mismas que ha elegido la naturaleza [6].

Curiosamente, la comprensión del mundo que nos rodea nos parece un desafío que logramos superar muy a menudo. La mente humana ha evolucionado inmersa en un entorno cuya complejidad está a la altura de nuestro intelecto en muchos aspectos (aunque no en todos). Encontramos patrones y los reconocemos de manera instintiva por lo valiosa que es esa capacidad para la supervivencia. El traslado de esa capacidad al mundo concreto y, después, al mundo abstracto de las matemáticas; el estudio de los patrones y, en última instancia, de las leyes de la naturaleza, comenzó con la sencilla observación de que $1 + 1 = 2$. Y sigue vigente.

ÍNDICE ANALÍTICO

[redacted]

A

B

C

D

E

F

G

H

I

J

K

L

M

N

O

P

Q

R

S

T

U

V

W

Y

Z

JOHN DAVID BARROW (Londres, 1952 - Cambridge, 2020) fue un físico, escritor, profesor universitario, divulgador científico, matemático, astrónomo y astrofísico británico.

Se doctoró en Oxford en 1977. Después colaboró con los departamentos de Física y Astrofísica de la Universidad de Oxford y de la Universidad de California, Berkeley. En 1999 pasó a ser catedrático de matemáticas y de física teórica en la Universidad de Cambridge. Ese mismo año obtuvo la Kelvin Medal de la Royal Philosophical Society of Glasgow. Ha sido uno de los más reconocidos astrofísicos de la actualidad. Entre sus numerosos títulos publicados cabe destacar *La trama oculta del universo: contar, pensar y existir* (1996), *¿Por qué el mundo es matemáticos?* (1997), *El libro de la nada* (2002), *Las constantes de la naturaleza* (2006) y *El salto del tigre* (2006) entre otros.

Notas

1. 1 + 1: ¿DE VERDAD ES ALGO TAN COMPLEJO?

[1] La estrofa viene a decir: «Uno es uno y así es y será por siempre jamás». (*N. de la T.*) <<

[2] La demostración comienza en la página 83 del volumen 2, donde aparece etiquetada como Proposición 110.643.22 <<

[3] Técnicamente pasa de ser un problema computacional de clase P a uno de clase NP. < <

2. LOS DEDOS DE LAS MANOS Y DE LOS PIES: CÓMO EMPEZAMOS A CONTAR

[1] Yo solía preguntar en ambientes matemáticos y al público general por qué habría de elegir alguien contar en base 8. Nadie me supo responder hasta que di una charla a niños de primaria (hasta 10 años de edad) y formulé la misma pregunta. Una niña me dio la respuesta correcta al instante y dijo que era algo obvio porque ella practicaba juegos con cuerdas entre los dedos, como el de la cuna del gato (que consiste en que dos personas formen figuras alternadamente con una cuerda entrelazada en los dedos de ambas manos). <<

[2] En la obra anterior de J. D. Barrow titulada *La trama oculta del universo* se ofrece un estudio más detallado de los sistemas antiguos de cómputo, que aparecen algo más resumidos en la obra de Barrow *¿Por qué el mundo es matemático?* y en la obra editada por G. Flegg titulada *Numbers Through the Ages* (Macmillan, 1989). <<

[3] «No hay dos cosas que puedan ser más iguales». < <

[4] R. Recorde, *The Urinal of Physick*, Londres, Brazen Serpent, 1547. < <

[5] Los logaritmos neperianos lograron reducir la multiplicación a sumas, y las divisiones a sustracciones. Para multiplicar A por B hay que escribir A y B como potencias de 10, de modo que $A = 10^a$ y $B = 10^b$, y $A \times B = 10^{a+b} = C$, por ejemplo. Las tablas de Napier indican qué valen a y b para A y B. Después, para calcular C se usan sus tablas a la inversa y se halla C. Por ejemplo, si $A = 2$ y $B = 3$, entonces $a = 0.3010$ y $b = 0.4771$, de modo que $C = 10^{0.7781} = 6.0$, tal como era de esperar. Así es como se realizaban todos los cálculos laboriosos antes de que se inventaran las calculadoras mecánicas y electrónicas. <<

[6] Para conocer más detalles consúltese el capítulo «Los tiempos de los signos» de la obra de J. D. Barrow titulada *Imágenes del Cosmos*, Barcelona, Paidós, 2009; traducción de Isabel Febrián y Cristina García, p. 261. < <

[1] En el caso de los babilonios, por motivos contables; en el caso de los mayas, por razones estéticas, para que no quedaran huecos en los glifos pictóricos con los que representaban los elementos constitutivos de los números grandes; y en el caso de los indios, por eficiencia computacional. < <

[2] El término *bit* lo acuñó por primera vez Claude Shannon en una publicación suya de 1948, pero se lo atribuyó a John W. Tukey, quien en 1947 había escrito un informe para los Laboratorios Bell en el que había empleado el vocablo *bit* como abreviatura de *binary information digit* («dígito de información binaria»). El *byte* es otra unidad de información digital que equivale a 8 bits, y se introdujo en la terminología computacional como el número de bits necesarios para especificar un solo carácter de texto. Es la unidad de memoria más pequeña a la que se puede acceder en la mayoría de arquitecturas computacionales, lo cual depende del *hardware* empleado. < <

[3] Para consultar una traducción al inglés de esta exposición de Leibniz, véase http://www.leibniz-translations.com/binary.htm <<

[4] J. W. Shirley, «Binary Enumeration before Leibniz», *American J. Physics*, 8, 452 (1951). < <

[5] En la acepción inicial del término, que en su origen aludía a una persona que realiza un cálculo de forma manual, no a una máquina calculadora. < <

[6] Para un sistema de n componentes, una descripción completa de su estado en la física clásica requiere tan solo n bits, mientras que en la física cuántica requiere 2n – 1 números complejos. Véase P. Shor, «Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer», *SIAM Journal on Computing*, 26, 1484 (1996), https://arxiv.org/abs/quant-ph/9508027. < <

[1] H. Bethe, citado en la revista *Life Magazine*, 1957, p. 89. Von Neumann era legendario por sus impresionantes dotes mentales y su velocidad de pensamiento: su capacidad para realizar cálculos matemáticos instantáneos, sus recuerdos inmediatos, su memoria fotográfica y su habilidad para traducir al instante entre varios idiomas. Se decía que trabajar con él era como ir en bicicleta persiguiendo un coche de carreras. Su compañero de colegio Eugene Wigner, galardonado con el premio Nobel, dijo: «Quien oye a Von Neumann entiende cómo debería funcionar la mente humana». Von Neumann realizó aportaciones fundamentales en muchos ámbitos de la lógica, la arquitectura computacional, las matemáticas, la física cuántica, la hidrodinámica de la física nuclear, la propagación de las ondas de choque, teoría de juegos, estadística y economía, y fue clave como asesor científico del gobierno estadounidense hasta su lamentadísima muerte a la temprana edad de 53 años. < <

[2] Hay teoremas que se pueden demostrar en el caso del ajedrez; por ejemplo, si las blancas tienen rey y reina, pero a las negras solo les queda el rey, las blancas no podrán hacer jaque mate a las negras. < <

[3] Por ejemplo, si los puntos *A, B, C* y *D* forman una línea recta tal que *B* se encuentra entre *A* y *C*, y *C* cae entre *B* y *D*, entonces no es posible *demostrar* a partir de los axiomas originales de Euclides que *B* se encuentra entre *A* y *D* (sino solo con un dibujo). < <

[4] H. C. Kennedy, *Peano. Life and Works of Giuseppe Peano*, Reidel (1980). < <

[5] G. Frege, versión en castellano: *Fundamentos de la aritmética: investigación lógico-matemática sobre el concepto de número*, trad. de Ulises Moulines, Barcelona, Laia, 1973; 2.ª edición. < <

[6] R. Dedekind, *Was sind und was sollen die Zahlen?*, Braunschweig, Vieweg, 1888. Versión en castellano: *¿Qué son y para qué sirven los números?, y otros escritos sobre los fundamentos de la matemática*, trad. de José Ferreirós Domínguez, Madrid, Alianza Editorial, 1998. < <

[7] La «hipótesis de la relatividad lingüística», también conocida como la hipótesis de Sapir-Whorf, aparece definida en el artículo titulado «The Status of Linguistics as a Science» (1956), del antropólogo y lingüista estadounidense E. Sapir, y también figura desarrollada en el artículo titulado «Science and Linguistics», de B. L. Whorf. < <

[8] Matemáticos y filósofos han luchado contra la idea de que tuvieran que existir las geometrías no euclídeas, mientras que para navegantes y artistas era, necesariamente, algo obvio. La obra *El matrimonio Arnolfini*, pintura al óleo del maestro holandés Jan van Eyck de 1434, es un ejemplo clásico de una escena reflejada en un espejo convexo. Un enfoque que parece no haberse adoptado nunca es el de imaginar todas las reglas y posibilidades de la geometría euclídea plasmadas sobre una página plana, como el teorema de Pitágoras, para verlas después reflejadas en un espejo curvo. Como es natural, todas las relaciones se mantienen y se da una correspondencia de uno a uno entre los puntos de la geometría plana y los del espejo curvo, de acuerdo con las leyes de la reflexión en espejos curvos que debemos a Fermat y otros que todos aprendimos en la asignatura de física del colegio. Por tanto, tenía que haber una base axiomática con coherencia interna para la geometría de las superficies curvas, y esta tenía que consistir en la transformación de la geometría de las superficies planas dictada por la ley de Fermat (1660) para la reflexión en espejos curvos (véase J. D. Barrow, *La trama oculta del universo*, Barcelona, Crítica, 1996; trad. de Javier García Sanz). Los espejos curvos hechos de vidrio o de metal pulido se han usado desde hace mucho tiempo, mientras que los espejos cóncavos se han utilizado para ver objetos distantes. < <

[1] B. Reznick, «A set is a set», en *Mathematics Magazine*, 66 (n.º 2), p. 95 (abril 1993). («Un conjunto es un conjunto; eso se da por seguro) y la nada no puede ser conjunto, ¡no hay duda! Pero eso será así, cariño, solo hasta que conozcas mi conjunto especialísimo»). < <

[2] El extraordinario matemático André Weil (hermano de la célebre filósofa Simone Weil) dice que introdujo el símbolo Ø tras tomarlo del alfabeto noruego; véase la autobiografía de André Weil: *Memorias de aprendizaje* (Tres Cantos, Nivola, 2002; trad. de Aurora Bell-Lloch). < <

[3] Cuando Bertrand Russell envió este ejemplo (ahora conocido como la «paradoja de Russell») a Frege, este último decidió abandonar su proyecto de encontrar una aritmética basada únicamente en la lógica. < <

[4] J. von Neumann, «Zur Einführung der transfiniten Zahlen», Acta litterarum ac scientiarum Regiae Universitatis Hungaricae FranciscoJosephinae, en *Sectio scientiarum mathematicarum*, 1, pp. 199-208 (1923); trad. al inglés *From Frege to Gödel: A Source Book in Mathematical Logic*, 1879-1931, de J. van Heijenoort, 3.ª edición, Cambridge, Massachusetts, Harvard University Press, pp. 346-354. < <

[5] El rey Lear, Acto 1, Escena 1. Esta frase forma parte de una tradición de paradojas nihilistas muy apreciadas por los escritores de este periodo porque les permitían decir cosas de manera ambigua que de otro modo se considerarían prosa herética, con lo que, en caso de que se les cuestionara, podían argumentar que en realidad estaban criticando la idea; véase el capítulo 2 de la obra *El libro de la nada* (Barcelona, Crítica, 2012; trad. de Javier García Sanz) de J. D. Barrow para consultar una exposición y ejemplos de esta tradición nihilista en la literatura. <<

[6] Esto es así porque en este caso hay que sumar cantidades que tienen una magnitud y una dirección (vectores). Si las fuerzas soportadas actuaran desde ángulos distintos a la perpendicular, entonces el resultado de la suma tampoco daría 1.414. < <

6. DEMOSTRACIÓN DE WHITEHEAD Y RUSSELL DE 1 + 1 = 2

[1] T. S. Eliot, «Commentary», *The Monthly Criterion*, núm. 6, octubre de 1927, p. 129. < <

[2] I. Grattan-Guinness, «The Royal Society's financial support of the publication of Whitehead and Russell's *Principia Mathematica*», en *Notes and Records of the Royal Society of London*, 30, 90, 1975. < <

[3] G. H. Hardy, *A Mathematician's Apology*, Cambridge University Press, 1940, p. 83; Versión en castellano: *Apología de un matemático*, Madrid, Capitán Swing Libros, 2017; trad. de Pedro Pacheco González. < <

[4] Stephen Wolfram señala que en el año 2000 realizó una búsqueda computacional del espacio de todos los posibles sistemas de axiomas, y descubrió el sistema más simple de axiomas para una lógica proposicional estándar. Y a partir de ese resultado pudo valorar en qué lugar reside esa lógica dentro del espacio de todos los sistemas axiomáticos formales posibles: dentro de una enumeración natural de sistemas de axiomas ordenados por tamaño, es el sistema formal que encontraríamos más o menos en el puesto número 50 000. La mayoría de los sistemas matemáticos que usamos necesita un sistema mayor que esta lógica básica; este artículo está disponible en la página web de Wolfram: **https://writings.stephenwolfram.com/2010/11/100-years-since-principia-mathematica/** <<

[5] Este teorema demuestra que cuatro es el número más bajo de colores que se puede usar para indicar los países de un mapa sin que ninguno de ellos porte el mismo color que alguno de los adyacentes. < <

[6] G. Boole, *An Investigation of the Laws of Thought*, Walton & Maberly, Londres, 1854. < <

[7] En el caso de los números naturales, la suma se corresponde con el conjunto de unión, y la multiplicación, con el producto cartesiano. En cambio, con la aritmética binaria, que solo dispone de 0 y 1, el conjunto de intersección se corresponde con la multiplicación, y el conjunto de unión equivale a la suma menos la multiplicación. < <

[1] En *The Saturday Review*, 15 de abril de 1978. < <

[2] También tuvieron gran profundidad las famosas cuatro paradojas de Zenón, formuladas por Zenón de Elea alrededor del año 450 a. C., y que no se resolvieron hasta el siglo XX. Aristóteles argumentó que la suposición de Zenón de que en un tiempo finito podía suceder una cantidad infinita de cosas era un error de razonamiento. <<

[3] R. Descartes, *Principles of Philosophy*, 27, citado en M. Blay, *Reasoning with the Infinite*, Univ. Chicago Press, Chicago, 1993. < <

[4] Para acceder a una exposición amplia de la historia de las distintas posturas frente a la noción de infinito, consúltese la obra de J. D. Barrow titulada *The Infinite Book*, J. Cape, Londres, 2005. <<

[5] San Agustín, *La ciudad de Dios*, Libro XII, capítulo 18. < <

[6] Como ejemplo sencillo, supongamos que lanzamos una moneda en vertical con una velocidad V desde una altura H por encima del suelo. El objeto ascenderá hasta una altura $h = H + Vt - 1/2gt2$ después de un tiempo t, donde g es la aceleración debida a la gravedad. Así que volverá a la mano del lanzador donde $h = H$ después de un tiempo $T = 2V/g$. Si al lanzarla al aire se le imprimió un giro de R revoluciones por segundo (*rps*), entonces habrá girado completamente N veces donde $N = T \times R = 2VR/g$. A partir de ahí se puede predecir el resultado de un lanzamiento que imprime al objeto N revoluciones. Si N vale 1, entonces al lanzar la moneda con la cruz hacia arriba, caerá con la cruz hacia arriba. Si N vale entre 2 y 3, 4 y 5, 6 y 7, o cualquier otro intervalo subsiguiente como estos, caerá del mismo lado con el que se lanzó. Pero si se lanzó con un valor de N entre 3 y 4, 5 y 6, 7 y 8, etc., entonces caerá con el lado contrario al del lanzamiento hacia arriba. A medida que N aumenta y pasa de 20, por ejemplo, las condiciones de V y R que determinan cada uno de los dos resultados distintos se acercan cada vez más, y solo se necesitan diferencias mínimas en las condiciones del lanzamiento para obtener como resultado cara o cruz. Por lo común V vale unos 2 m/s; mientras que $g = 9.8\ m/s2$; por tanto, $2V/g = 0.4\ s$. Para que dé tiempo a que la moneda gire más de 20 veces y garantizar un resultado cercano a
50-50,
la moneda debe lanzarse de manera que el ritmo de rotación supere $20/0.4 = 50\ r/s$. Para más detalles véase J. D. Barrow, *100 Things you Didn't Know About Maths and Sport*, Bodley Head, Londres, 2012, capítulo 94. <<

[7] G. Galilei, Consideraciones y demostraciones matemáticas sobre dos nuevas ciencias, Madrid, Editora Nacional, 1981; trad. de Javier Sádaba Garay. < <

[8] G. Galilei, Diálogos acerca de dos nuevas ciencias, Buenos Aires, Losada, 1945; trad. al castellano de José San Román Villasante. <<

[9] Segunda carta de Newton a Bentley sobre el infinito, véase I. B. Cohen, *Isaac Newton's Papers & Letters on Natural Philosophy*, Cambridge, Mass., Harvard University Press, 1958. Versión en castellano extraída de Isaac Newton, *Cuatro cartas al Dr. Bentley. Carta al honorable Sr. Boyle sobre la causa de la gravitación*, Madrid, Editorial Complutense, 2001; trad. de Luis Rodríguez Luján y José Luis González Recio. < <

[10] Cabe preguntarse en qué lugar de esta jerarquía se sitúa el continuo. Cantor sostuvo que constituía el primer paso después de $\aleph_0$ y, por lo tanto, se correspondía con $\aleph_1$. Su conjetura recibió el nombre de la hipótesis del continuo, y su solución se convirtió en uno de los problemas más complejos de la teoría de conjuntos. Kurt Gödel en 1938 y Paul Cohen en 1963 revelaron que el problema no podía demostrarse ni refutarse usando los axiomas estándar de la teoría de conjuntos: es indecidible. <<

[11] **https://www.uni-siegen.de/fb6/phima/lehre/phima10/quellentexte/handout-phima-teil4b.pdf.** < <

[12] Esto sugiere que si se sigue el revolucionario razonamiento de Cantor hacia atrás podría llegarse a un infinito absoluto negativo ¡que podría usarse para explicar la existencia del diablo! Pero Cantor nunca pensó en infinitos negativos, que en cualquier caso no encajan bien en su construcción. En cuanto a los infinitesimales, Cantor se opuso firmemente a su uso en matemáticas. < <

[13] J. W. Dauben, en *Journal of the History of Ideas*, 38, 89, 1977, núm. 15. < <

[14] 14. D. Burton, *History of Mathematics*, Dubuque, Ia., W.C. Brown, 1995, 3.º de., p. 593. < <

[15] Ampliada y modificada a partir de R. Rucker, *Infinity and the Mind*, Princeton, Nueva Jersey, Princeton University Press, 1995. < <

[16] D. Hilbert, «Über das Unendliche», en *Mathematische Annalen*, 95, 1, 1926, pp. 161-190, cita en p. 170. < <

[1] J. D. Barrow, *Impossibility*, Oxford UP, Oxford, (1998). Versión en castellano: *Imposibilidad: los límites de la ciencia y la ciencia de los límites*, Barcelona,: Gedisa, 1999; trad. de Carlos de la Reta. < <

[2] J. D. Barrow, The Book of Nothing, Cape, Londres, 2000. Versión en castellano: El libro de la nada, Barcelona, Crítica, 2012; trad. de Javier García Sanz. < <

[3] E. Grant, *Much Ado About Nothing: Theories of Space and Vacuum from the Middle Ages to the Scientific Revolution*, Cambridge, Cambridge University Press, 1981. Hay ejemplos previos de algunos de estos provocadores experimentos mentales en el Libro I de la obra *De rerum natura* de Lucrecio. < <

[4] La obra H. Leff y A. Rex (eds.), *Maxwell's Demon*, Princeton, Princeton University Press, 1990, ofrece una visión general con la reimpresión de los artículos científicos más relevantes. < <

[5] K. Guthrie (ed.), *The Pythagorean Sourcebook and Library*, Phares Press, Grand Rapids, (1987). < <

[6] P. Pesic, *Abel's Proof: An Essay on the Sources and Meaning of Mathematical Unsolvability*, Cambridge, Mass, MIT Press, 2003. Este libro también contiene una nueva traducción del artículo de Abel de 1824 con su hallazgo. < <

[7] La fórmula para las de segundo grado la conocemos de la escuela: $Ax2 + Bx + C = 0$, la cual tiene soluciones en términos de A, B y C que vienen dadas por la fórmula $x = [–B \pm \sqrt{(B2 – 4AC)}]/2A$. < <

[8] Véase J. Gray, The Hilbert Challenge, Oxford, Oxford University Press, 2000, que contiene el texto de la conferencia de Hilbert en el Congreso Internacional de Matemáticas de 1900, pp. 240-282; véase también M. Toepell, Archive for History of Exact Sciences, 35, 329, 1986. < <

[9] D. Hilbert, «Über das Unendliche», en Mathematische Annalen, 95, 1, 1926, pp. 161-190, p. 170. < <

[10] L. Corry, en *Archive for History of Exact Sciences*, 51, 83, 1997. < <

[11] C. Reid y H. Weyl, *Hilbert*, Berlín, Springer Verlag, 1970, p. 92.
< <

[12] J. R. Lucas en *Philosophy* 36, 112, 1961 y *Freedom of the Will*, Oxford, Clarendon, 1970. Para acceder a una exposición de calidad de la demostración de Gödel y las ramificaciones más amplias de su teorema véase T. Franzén, *Gödel's Theorem: an incomplete guide to its use and abuse*, Wellesley, Massachusetts, A. K. Peters, 2005. <<

[13] S. Jaki, *Cosmos and Creator*, Edimburgo, Scottish Academic Press, 1980, pp. 4955 e Id., *The Relevance of Physics*, Chicago, The University of Chicago Press, 1970. < <

[14] J. Dryden, *Absalom and Achitophel*, Oxford, Clarendon Press, 1911. < <

[15] A. Tarski, A. Mostowski y R. M. Robinson, *Undecidable Theories*, Ámsterdam, North Holland, 1953. < <

9. POR QUÉ SON TAN COMUNES LOS UNOS Y LOS DOSES

[1] Bobby Robson fue un futbolista británico de talla internacional, seleccionador del equipo nacional de Inglaterra y entrenador del Newcastle United. Citado en la página de deportes de *The Independent*, 2 de febrero de 2016 en un artículo de Chris McGrath. <<

[2] S. Newcomb, en *American Journal of Mathematics*, 4, 39, 1881.
< <

[3] Esto se aplica a cantidades que tienen dimensiones, como las áreas. De hecho, la distribución que encontraron Newcomb y Benford tiene la propiedad de permanecer igual al reescalar las cantidades para expresarlas en otras unidades. Esta condición invariable con la que la distribución del primer dígito responde a $P(kx) = \int(k)P(x)$, donde k es una constante, y $\int(k)$ es una función de k, requiere que $P(x) = 1/\text{x}$ y que $\int(k) = 1/k$ y selecciona la distribución de Newcomb-Benford únicamente como $P(d) = [\int d\text{d} +1\ dx/x]/\ [\int 110\ dx/x] = \log10[1 + 1/d]$. <<

[4] M. Nigrini, *The Detection of Income Evasion through an Analysis of Digital Distributions*, tesis doctoral, Universidad de Cincinnati, 1992. < <

[5] Su ley se deriva con exactitud de cualquier proceso que tenga una probabilidad de distribución $P(x) = 1/x$ siempre que x caiga en el intervalo de 0 a 1. Si esta distribución de probabilidad fuera $P(x) = 1/x$ con una forma diferente de 1, entonces los resultados serían distintos, y la probabilidad de que el primer dígito sea d pasa a ser $P(d) =$
(101-a
– 1)-1[(d + 1)1-a – d1-a]. Para a = 2, $P(1)$ vale ahora 0.56. <<

10. ¿QUÉ SON LAS MATEMÁTICAS?

[1] Véase **https://www.math.utah.edu/~cherk/mathjokes.html** < <

[2] Øystein Linnebo ofrece una buena panorámica de todas las posibilidades y problemas en **https://plato.stanford.edu/entries/platonismmathematics/** < <

[3] P. Ernest, *Social Constructivism as a Philosophy of Mathematics*, Albany, State University of New York Press, 1998; R. Hersh, *What is Mathematics Really?*, Londres, Vintage, 1998. <<

[4] El mundo podría seguir siendo matemático aunque casi todos sus teoremas fueran incomputables en el sentido de Turing. Solo tendríamos demostraciones y teoremas no constructivistas que no serían de utilidad para producir algoritmos que calculen lo que vemos. < <

[5] E. Wigner, «The Unreasonable Effectiveness of Mathematics in the Natural Sciences», en *Communications in Pure and Applied Mathematics*, 13, 1, 1960. A esta afirmación se le podría añadir otra que hiciera hincapié en la irrazonable ineficacia de las matemáticas en las ciencias sociales y políticas. <<

[6] P. A. M. Dirac, en Proceedings of the Royal Society of Edinburgh, 59, parte 2, 124, 1938-1939. < <